

PCT/GB01/00299

WO 01/55386

Toll-like Receptor

Field of the Invention

The present invention relates to a novel Toll-like receptor or a variant thereof. A variant may demonstrate Toll-like receptor activity such as activation of NFκB, or may inhibit Toll-like receptor activity.

Background of the Invention

A family of human Toll-like receptors has been described in the literature. These receptors are termed Toll-like receptors in view of common structural features shared with a Drosophila Toll (dToll) receptor molecule which is involved in embryonic development. Toll and Toll-like receptors are type I transmembrane proteins, with extracellular leucine rich repeat motifs and an intracellular signalling domain homologous to that of members of the interleukin 1 receptor superfamily.

Drosophila Toll also plays an important role in the adult fly and is involved in immune surveillance mechanisms required for recognition of bacterial and fungal pathogens and regulation of specific innate immune defence gene expression. Activation of dToll receptors in response to infection by specific micro-organisms is thought to require the production of a protein ligand called Spaetzle. The human Toll-like receptors (hTLRs) are also thought to participate in mechanisms of innate immunity and inflammation acting as pattern recognition receptors (PRRs) for bacteria and other micro-organisms. hTLRs are expressed on antigen presenting cells including monocytes and dendritic cells and like dToll play roles in innate immunity. TLRs can elicit pro-inflammatory cytokine production and induce expression of cell surface co-stimulatory receptors required for activation of T-cells. Some hTLRs may help to co-ordinate interactions between cells of the innate and acquired immune systems to orchestrate an integrated immune response to infection.

(13) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2001 (02.08.2001)

(10) International Publication Number
WO 01/55386 A1

(31) International Patent Classification: C12N 15/12, 15/63, 5/10, G01N 33/68, C07K 14/703, 16/28, A61K 31/7083, 39/17, 48/00

(32) International Application Number: PCT/GB01/00299

(33) International Filing Date: 25 January 2001 (25.01.2001)

(34) Filing Language: English

(35) Publication Language: English

(36) Priority Date: 23 January 2000 (23.01.2000)

(37) Applicant (for all designated States except US): GLAXO-SMITHKLINE (GB/GB); Corporate Intellectual Property, Two New Horizons court, Brentford, Middlesex TW8 9BP (GB).

(38) Inventor(s) and
(39) Invention/Applicant (for US only): LEWIS, Alan, Peter (GB/GB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB); RAY, Keith, Paul (GB/GB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB).

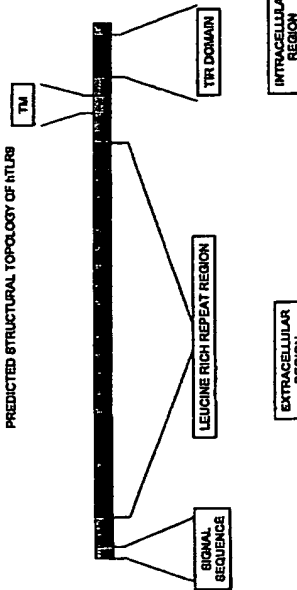
(40) Title: TOLL-LIKE RECEPTOR

with international search report
before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guide to the Preparation of Claims and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: TOLL-LIKE RECEPTOR

PREDICTED STRUCTURAL TOPOLOGY OF hTLR



(57) Abstract: An isolated Toll-like receptor polypeptide comprises the amino acid sequence of SEQ ID NO: 2, a variant or a fragment thereof which has immunomodulatory activity. Polynucleotides encoding such a Toll-like receptor are also described. A method for identification of a substance that modulates Toll-like receptor activity comprises contacting a polypeptide of the invention with a test substance and monitoring for immunomodulatory activity.

The specific functions of two mammalian TLRs, TLR2 and TLR4, have recently been identified. TLR2 and TLR4 are involved in mediating host responses to gram positive and gram negative bacteria through recognition of specific bacterial wall components. It has also recently been shown that TLR4 mediates responses to certain viral proteins including respiratory syncytial virus (RSV) (Nature Immunology 1: 398 2000).

Additionally TLRs may form heterodimeric functional complexes. Components of the intracellular signal transduction pathways of some hTLRs appear to be shared with Interleukin-1 (IL-1) receptor transduction pathways. Stimulation of TLR2 and TLR4 leads to activation of NF- κ B via an adaptor protein MyD88 and recruitment of the IL-1 receptor associated kinases (IRAKs).

Summary of the invention

A novel Toll-like receptor is now provided which is a screening target for the identification and development of novel pharmaceutical agents which modulate the activity of the receptor and in particular have immunomodulatory activity. These agents may be used in the treatment and/or prophylaxis of inflammatory diseases, cardiovascular diseases, systemic infections and autoimmune diseases, such as asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, viral, fungal and bacterial infections, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis. These agents may also be used as immunoadjuvants to enhance or alter the immune response in vaccine therapy.

Accordingly, the present invention provides an isolated Toll-like-receptor polypeptide which comprises:

- (i) the amino acid sequence of SEQ ID NO: 2;
- (ii) a variant of (i) which has immunomodulatory activity; or
- (iii) a fragment of (i) or (ii) which retains immunomodulatory activity.

Preferably, a variant has at least 80% identity to the amino acid sequence of SEQ ID NO: 2, more preferably at least 85% identity therewith, for example 87% identity therewith.

The invention also provides a polynucleotide encoding a polypeptide of the invention. Such a polynucleotide may be a polynucleotide which encodes a Toll-like receptor polypeptide which has immunomodulatory activity. The polynucleotides of the invention may be DNA or RNA, for example mRNA. A polynucleotide according to the invention comprises:

- (a) the nucleic acid sequence of SEQ ID NO: 1 and/or a sequence complementary thereto;
- (b) a sequence which hybridises under stringent conditions to a sequence as defined in (a);
- (c) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a) or (b); or
- (d) a sequence having at least 80% percent identity to a sequence as defined in (a), (b) or (c).

The present invention also provides a polypeptide expressed from a polynucleotide according to (a), (b), (c) or (d) above, in particular a polypeptide comprising a toll-like receptor according to the invention, encoded by the mRNA derived from a DNA sequence according to (a) or (b) above, thus the invention provides an isolated toll-like receptor polypeptide which is obtainable by expression *in vitro* or *in vivo* of a DNA molecule comprising the sequence of nucleotides as shown in SEQ ID NO.1.

The polypeptides of SEQ ID NO 2 and SEQ ID NO 4, which are different isoforms expressed from the multiple exon *tlr9* gene, are herein referred to as TLR9 and TLR9-A, respectively. TLR9-A is encoded by the nucleotide sequence of SEQ ID No. 3, which is encoded within SEQ ID No.1 except for the Initiating methionine, that is encoded by a second exon as illustrated in figure 1. (see Hemmi et al. Nature 408, 740-745 2000; Du et al, European Cytokine Network, 11: 382-371, 2000; Chuang and Ulevitch, European Cytokine Network, 11: 372-378 for isolation of the cDNAs, and corresponding sequence database accessions EMBL:AB045180, EMBL:AF259282, EMBL:AF245704).

10

In further aspects of the invention we provide:

- an expression vector capable of expressing a polypeptide of the invention comprising a polynucleotide as defined above.
- a host cell comprising an expression vector of the invention.
- an antibody specific for a polypeptide of the invention.
- a method for identification of a compound that modulates Toll-like receptor activity, comprising contacting a polypeptide of the invention with a test compound and monitoring for immunomodulatory activity.

15

20. Compounds which are identifiable in accordance with this method may be used in the treatment of a subject having a disorder that is responsive to Toll-like receptor modulation such as an inflammatory or cardiovascular disorder or systemic infection or autoimmune disease, including asthma, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, viral, fungal and bacterial infections, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes and Multiple Sclerosis. In particular, compounds which are identifiable in accordance with this method may be used in the treatment of a subject having allergic asthma or

30

rinitis. Further, such compounds may have immunomodulatory activity and be of use in the treatment of, or as adjuvants in vaccination against, bacterial or viral infections or as components of anti-cancer vaccines.

5 Compounds identifiable in accordance with this method include, in particular, synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which bind directly to the Toll-like receptor of the invention, and synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

10

15 In an alternative aspect of the invention, a polypeptide comprises a fragment or variant of SEQ ID NO 2 which is capable of inhibiting the activity of TLR9 or TLR9-A, for use in the treatment of an immune or inflammatory disorder.

In a further aspect of the invention, a polypeptide or polynucleotide in accordance with the invention or a compound identifiable in accordance with the invention is provided for use as an adjuvant or as an immunotherapeutic agent, for example in a vaccine.

20

Brief Description of the Sequences

SEQ ID NO: 1 is the amino acid sequence of human protein TLR9 and its encoding DNA;

25

SEQ ID NO: 2 is the amino acid sequence alone of TLR9;

SEQ ID NO: 3 is the amino acid sequence of human protein TLR9-A and its encoding cDNA (EMBL:AF259282);

SEQ ID NO: 4 is the amino acid sequence alone of TLR9-A (Hemmi et al.).

30

Brief Description of the Drawings

Fig. 1 is a diagrammatic illustration showing the exon arrangement encoding TLR9 and TLR9-A;

Fig. 2 shows tissue distribution data for TLR9 (using a human tissue cDNA plate). The profile shows predominant expression in tonsil and adenoid tissues with lower levels of expression detected in adipose, adrenal, foetal brain, cerebellum, jejunum, lung, myometrium, omentum, head of pancreas, rectum, skeletal muscle, spleen and thymus tissues;

Fig. 3 shows tissue distribution data for TLR9 (using a human disease cDNA plate). The profile shows predominant expression in lung tissue, bone marrow and PBMC with lower levels of expression detected in some colon, breast and brain/cerebellum samples;

Fig. 4 illustrates, in diagrammatic form, the predicted structural topology of human TLR9 - "TM" is the transmembrane portion, "TIR" is the cytosolic region conserved among interleukin and toll-like receptors known as the Toll Interleukin Receptor domain.

Detailed Description of the Invention

A single open reading frame was identified in genomic DNA (SEQ ID No.1), which encodes a protein of 1055 residues, predicted molecular weight 118,515 (PeptideSort - GCG Software) and the amino acid sequence shown in SEQ ID No.2. This sequence included the TIR domain common to Toll-like receptors and members of the interleukin-1 receptor family e.g. IL1RI, and the N-terminal sequence contains structural features as shown in figure 4. These features include, in order from the N- to the C-terminus, a predicted signal sequence with a potential cleavage site between residues 48 and 49 or 50 and 51 (SPScan in GCG; SignalP), a leucine-rich repeat motif domain, a potential transmembrane region and the Toll/IL-1R homologous region (TIR; Pfam Database). These motifs confirmed that the protein was likely to be expressed as a single transmembrane receptor-like molecule belonging to the TLR rather than the IL1R

family and therefore, it was designated TLR9 based on existing published and in-house nomenclature.

Throughout the present specification and the accompanying claims the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

The present invention relates to a human Toll-like receptor, referred to herein as TLR9, and variants or fragments thereof. Sequence information for TLR9 is provided in SEQ ID NO: 1 (nucleotide and amino acid) and in SEQ ID NO: 2 (amino acid only). A polypeptide of the invention consists essentially of the amino acid sequence of SEQ ID NO: 2 or of a functional variant of that sequence. One important variant of TLR9 is TLR9-A, sequence information for which is provided in SEQ ID NO: 3 (nucleotide and amino acid) and in SEQ ID NO: 4 (amino acid only).

The polypeptides are provided in isolated form. The term "isolated" is intended to convey that the polypeptide is not in its native state, insofar as it has been purified at least to some extent or has been synthetically produced, for example by recombinant methods. The polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. The term "isolated" therefore includes the possibility of the polypeptide being in combination with other biological or non-biological material, such as cells, suspensions of cells or cell fragments, proteins, peptides, expression vectors, organic or inorganic solvents, or other materials where appropriate, but excludes the situation where the polypeptide is in a state as found in nature.

A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the invention. Routine methods, can be employed to purify and/or synthesise the proteins according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Edition, CSH Laboratory Press (1989), the disclosure of which is included herein in its entirety by way of reference.

10

The term "variants" refers to polypeptides which have the same essential character or basic biological functionality as TLR9. The essential character of TLR9 can be defined as that of a Toll-like receptor. In particular, it refers to a polypeptide which has an immunomodulatory activity. In one aspect of the invention, a polypeptide of the invention may activate NF κ B or may elicit pro-inflammatory cytokine production or induce expression of cell surface co-stimulatory receptors required for activation of T-cells.

15

Alternatively, a variant of the polypeptide of the invention is one which exhibits binding to the same ligand as TLR9. Such ligand binding may be assayed using the assays described below.

20

In other aspects of the invention a variant is one which does not show the same function as TLR9 but which may be used to inhibit this function. For example, a variant polypeptide for use in an assay or therapy is one which inhibits TLR9 activity, for example by inhibiting or competing out ligand binding or receptor complex formation by TLR9. Alternatively, a variant may be one which inhibits ligand binding to TLR9. Such a variant may inhibit activation of NF κ B or inhibit cytokine production and expression of cell surface co-stimulatory receptors.

25

Such inhibitors may be used as immunomodulators to reduce inappropriate TLR activation in asthma or other chronic inflammatory diseases, or septic shock.

To determine whether a variant has the same essential function as TLR9, the immunomodulatory activity can be determined by monitoring the effect of a substance on different immune responses. For example the effect of the substance under test on NF κ B activation mediated through binding the polypeptide of the present invention may be monitored. This can be carried out, for example, by co-transfection of a construct expressing the polypeptide with a construct containing a reporter gene, such as secreted placental alkaline phosphatase, under the control of a suitable NF κ B-responsive promoter and monitoring for expression of the reporter gene.

10

Alternatively, other immunomodulatory activity such as the production of cytokines can be determined by monitoring cytokine production following incubation of a test substance with a cell expressing a polypeptide of the invention. Such assays may be carried out in the presence or absence of additional T-lymphocytes to assess the effect of such cytokines, or the direct action of a polypeptide of the invention, on such T-lymphocytes to thus determine immunomodulatory activity.

15

20

Alternatively, the Toll-like receptor functionality is as a peptide which binds a ligand of TLR9, inhibits immunomodulatory activity by TLR9 or inhibits ligand binding to TLR9 and can be determined by an assay as described below.

25

Preferably, a polypeptide of the invention will show the structural features associated with a Toll-like receptor. Preferably, a polypeptide of the invention, or a functional fragment thereof, contains one or more of the following structural features associated with a Toll-like receptor: an extracellular region containing leucine-rich repeat motif and cysteine-rich regions involved in ligand binding; a

30

single hydrophobic transmembrane region; and a C-terminal cytoplasmic domain sharing homology with other TLRs and members of the IL-1 receptor family.

Typically, polypeptides with more than about 65% identity, preferably at least 80% or at least 90% and particularly preferably at least 95%, at least 97%, or at least 99% identity, with the amino acid sequences of SEQ ID NO: 2 over a region of at least 20, preferably at least 30, at least 40, at least 60 or at least 100 contiguous amino acids or over the full length of SEQ ID NO: 2, are considered as variants of the proteins. Identity is calculated using the widely used GCG (University of Wisconsin) suite of programs and preferably using the distances software (correction method). Such variants may include allelic variants and the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains the basic biological functionality of the Toll-like receptor, having a similar function to TLR9 or inhibits such function such as preventing ligand binding or TLR9 mediated activation. Such variants also include isoforms such as TLR9-A, which is 23 amino acids (or 2.2%) shorter than TLR9 (see SEQ ID NO: 2 and SEQ ID NO: 4) and thus shows 97.8 identity therewith. Transcription of the nucleotide sequence presented in SEQ ID NO:1 can result, due to variable mRNA splicing involving a second exon encoding an alternative initiating methionine, in an mRNA having the sequence of SEQ ID NO: 3 which, when translated, results in the polypeptide TLR9-A depicted in SEQ ID NO: 4.

Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions. The modified polypeptide generally retains activity as a TLR9 receptor or inhibitor of TLR9 receptor activity. Conservative substitutions may be made, for example according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

ALIPHATIC	Non-polar	G A P
		I L V
	Polar-uncharged	C S T M
	Polar-charged	N Q D E K R H F W Y
AROMATIC		

Shorter polypeptide sequences are within the scope of the invention. For example, a peptide of at least 20 amino acids or up to 50, 60, 70, 80, 100 or 150 amino acids in length is considered to fall within the scope of the invention as long as it demonstrates the basic biological functionality of TLR9 or inhibits TLR9. In accordance with this aspect of the invention the peptide may also comprise an epitope of TLR9 for generation of antibodies. In particular, but not exclusively, this aspect of the invention encompasses the situation when the protein is a fragment of the complete protein sequence and may represent a ligand-binding region (N-terminal extracellular domain) or an effector binding region (C-terminal intracellular domain). Fragments from which the C-terminus has been removed may be used as decoy receptors. Other fragments such as a secreted or soluble form of the receptor may be generated for use in an assay or in therapy in accordance with the invention. Such fragments can also be used to raise anti-TLR9 antibodies.

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or may comprise modified amino acid residues. They may also be modified by the addition of histidine residues or an epitope tag for example by a (His) 6 or (His) 6 sequence or a HA, 77, Myc or Flag tag to assist their purification or detection. They may be modified by the addition of a signal sequence to promote insertion into the

cell membrane. Such modified polypeptides fall within the scope of the term "polypeptide" of the invention.

6 The invention also includes nucleotide sequences that encode for TLR9 or variants thereof as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA including genomic DNA, synthetic DNA or cDNA. Preferably the nucleotide sequence is a DNA sequence and most preferably, a cDNA sequence. Nucleotide sequence information is provided in SEQ ID NO: 1. Such nucleotides can be isolated from human cells or synthesised according to methods well known in the art, as described by way of example in Sambrook *et al.* Such nucleotides can typically be isolated from activated cells of the immune system, heart, lung, pancreatic islet cells and lymph nodes, adenoid and tonsil tissues. Figures 2 and 3 show the tissue distribution of RNA encoding TLR9, as determined by extraction of total RNA from normal or disease tissue or cells which is then used to generate cDNA for real time quantitative PCR using suitable primers and probes (TaqMan analysis) to assess expression patterns. The profiles show differential expression across tissues tested and predominance to sites containing inflammatory cell types.

20 Typically a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridising under selective conditions to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1.

25 A polynucleotide of the invention can hybridize to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1 (or of SEQ ID NO: 3) at a level significantly above background. Background hybridisation may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the coding sequence or complement of the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3 is typically at least 10 fold, preferably at least 100 fold, as intense

as interactions between other polynucleotides and the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ³²P. Selective hybridisation may typically be achieved using conditions of low stringency (0.03M sodium chloride and 0.03M sodium citrate at about 40°C), medium stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 50°C) or high stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 60°C).

10 The coding sequence of SEQ ID NO: 1 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 25, 50 or 100 substitutions. The polynucleotides of the present invention may alternatively or additionally be modified by one or more insertions and/or deletions and/or by an extension at either or both ends. The modified polynucleotide generally encodes a polypeptide which has Toll-like receptor activity or inhibits the activity of TLR9. Degenerate substitutions may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in the Table above.

20 A nucleotide sequence of the invention which is capable of selectively hybridising to the complement of the DNA coding sequence of SEQ ID NO: 1 will generally have at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, more preferably at least 100 contiguous nucleotides or most preferably over the full length of SEQ ID NO: 1. Methods of measuring nucleic acid and protein homology are well known in the art. For example the UWGCG Package provides the BESTFIT program which can be used to calculate homology (Devereux *et al* 1984). Similarly the PILEUP and BLAST algorithms can be used to line up sequences (for example are described in Altschul 1983,

and Altshul *et al* 1990). Many different settings are possible for such programs. In accordance with the invention, the default settings may be used.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher sequence identity over longer lengths) being preferred. Thus, for example a polynucleotide which has at least 80% sequence identity over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which has at least 95% sequence identity over 40 nucleotides. The most preferred sequences have at least 70% sequence identity over at least 70% of the full length of the sequence provided by SEQ ID NO: 1.

The nucleotides according to the invention have utility in production of the proteins according to the invention, which may take place *in vitro*, *in vivo* or *ex vivo*. The nucleotides may be involved in recombinant protein synthesis or indeed as therapeutic agents in their own right, utilised in gene therapy techniques. Nucleotides complementary to those encoding TLR9, or antisense sequences, may also be used in gene therapy, such as in strategies for down regulation of expression of the proteins of the invention.

Polynucleotides of the invention may be used as a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will preferably be at least 10, preferably at least 15 or at least 20, for example at least 25, at least 30 or at least 40 nucleotides in length. They will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150

nucleotides in length, for example up to 200, 300, 400, 500 nucleotides in length, or even up to a few nucleotides, such as five or ten nucleotides, short of the coding sequence of SEQ ID NO: 1.

The present invention also includes expression vectors that comprise nucleotide sequences encoding the proteins or variants thereof of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to a person skilled in the art. By way of further example in this regard we refer to Sambrook *et al*.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used as test compounds in the assays of the invention or may be useful in a method of treatment of the human or animal body by therapy.

Preferably, a polynucleotide of the invention or for use in the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence, such as a promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example in a method of gene therapy.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which expression is designed. For example, yeast promoters include *S. cerevisiae* GAL4 and ADH promoters, *S. pombe* *nmt1* and *adh* promoter. Mammalian promoters include the metallothionein promoter which can be induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

Mammalian promoters, such as β -actin promoters, may be used. Tissue-specific promoters may be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). Viral promoters are readily available in the art.

The vector may further include sequences flanking the polynucleotide which comprise sequences homologous to eukaryotic genomic sequences, preferably mammalian genomic sequences, or viral genomic sequences. This will allow the

Introduction of the polynucleotides of the invention into the genome of eukaryotic cells or viruses by homologous recombination. In particular, a plasmid vector comprising the expression cassette flanked by viral sequences can be used to prepare a viral vector suitable for delivering the polynucleotides of the invention to a mammalian cell. Other examples of suitable viral vectors include herpes simplex viral vectors and retroviruses, including lentiviruses, adenoviruses, adeno-associated viruses and HPV viruses (such as HPV-18 or HPV-16). Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for example may be used to stably integrate the polynucleotide giving rise to the RNA into the host genome. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression.

The invention also includes cells that have been modified to express the TLR9 polypeptide or a variant thereof. Such cells include transient, or preferably stable higher eukaryotic cell lines, such as mammalian cells or insect cells, lower eukaryotic cells, such as yeast or prokaryotic cells such as bacterial cells. Particular examples of cells which may be modified by insertion of vectors encoding for a polypeptide according to the invention include mammalian HEK293T, CHO, HeLa and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of a polypeptide. Cells such as T-cells, monocytes or dendritic cells expressing the receptor may be used for example in screening. Expression may be achieved in transformed oocytes. A polypeptide of the invention may be expressed in cells of a transgenic non-human animal, preferably a mouse. A transgenic non-human animal expressing a polypeptide of the invention is included within the scope of the invention.

It is also possible for the proteins of the invention to be transiently expressed in a cell line or on a membrane, such as for example in a baculovirus expression

18

system. Such systems, which are adapted to express the proteins according to the invention, are also included within the scope of the present invention.

5 According to another aspect, the present invention also relates to antibodies (either polyclonal or preferably monoclonal antibodies, chimeric, single chain, Fab fragments) which have been raised by standard techniques and are specific for a polypeptide of the invention. Such antibodies could for example, be useful in purification, isolation or screening methods involving immunoprecipitation techniques and may be used as tools to further elucidate the function of TLR9 or a variant thereof, or indeed as therapeutic agents in their own right. Antibodies may also be raised against specific epitopes of the proteins according to the invention. Such antibodies may be used to block ligand binding to the receptor. Alternatively an antibody may be provided which acts as an agonist, to cross link receptors of the invention to mediate receptor activity. An antibody, or other compound, "specifically binds" to a protein when it binds with high affinity to the protein for which it is specific but does not bind or binds with only low affinity to other proteins. A variety of protocols for competitive binding or immunoradiometric assays to determine the specific binding capability of an antibody are well known in the art (see for example Maddox et al 1993). Such immunoassays typically involve the formation of complexes between the "specific protein" and its antibody and the measurement of complex formation.

25 An important aspect of the present invention is the use of polypeptides according to the invention in screening methods to identify compounds that may act as modulators of Toll-like receptor activity. Any suitable form may be used for the assay to identify a modulator of TLR9 activity. In general terms, such screening methods may involve contacting a polypeptide of the invention with a test compound and then measuring receptor activity.

19

Screening methods may alternatively involve contacting a polypeptide of the invention with a test compound and then monitoring for the effect on immunomodulatory activity.

5 The binding of the substance to a polypeptide in the invention can be determined directly. For example, a radiolabelled test substance can be incubated with a polypeptide of the invention and so that binding of the test substance to the polypeptide can be monitored. Typically, the radiolabelled test substance can be incubated with cell membranes containing the polypeptide until equilibrium is reached. The membranes can then be separated from a non-bound test substance and dissolved in scintillation fluid to allow the radioactive content to be determined by scintillation counting. Non-specific binding of the test substance may also be determined by repeating the experiments in the presence of a saturating concentration of a non-radioactive ligand. Preferably, a binding curve is constructed by repeating the experiment with various concentrations of the test substance.

15 Cell based assays may also be carried out, for example using a cell expressing the TLR9 receptor, and contacting the cell with another cell to look for ligand binding or activation of TLR9-mediated pathways such as NF κ B activation. Alternatively cells expressing TLR9 constitutively may be provided for use in assays for TLR9 function. Such constitutively expressed TLR9 may demonstrate TLR9 activity in the absence of ligand binding. Additional test substances may be introduced in any assay to look for inhibitors of ligand binding or inhibitors of TLR9-mediated activity. Assays are preferably carried out using cells expressing TLR9, and incubating such cells with the test substance optionally in the presence of TLR9 ligand. Alternatively an antibody may be used to complex TLR9 and thus mediate TLR9-activity. Test substances may then be added to assess the effect on such activity.

20

In preferred aspects, a host cell is provided expressing the receptor, or a receptor complex of TLR9 (or TLR9-A) comprising a homodimer, a heterodimer of TLR8 (or TLR9-A) with another Toll-like receptor, or a complex of TLR9 (or TLR9-A) with protein cofactors, and containing an NFkB responsive reporter construct. The host cell is treated with a substance under test for a defined time. The expression of the reporter gene, such as secreted placental (SP) alkaline phosphatase or luciferase is assayed. The assay enables determination of whether the addition of compounds inhibits the induction of the response in target cells. Alternatively the assay may be carried out to identify cytokine production or it may be carried out in the presence of T-cells to identify induction of co-stimulatory receptors required for activation of T-cells.

10

Assays may also be carried out to identify modulators of receptor shedding. A polypeptide of the invention can be cleaved from the cell surface. Shedding the receptor would act to down regulate receptor signalling. Thus, cell based assays may be used to screen for compounds which promote or inhibit receptor-shedding. Assays may also be carried out to identify substances which modify TLR8 receptor expression for example substances which down regulate expression. Such assays may be carried out for example by using antibodies for TLR8 to monitor levels of TLR8 expression.

15

20

Additional control experiments may be carried out. Assays may also be carried out using known ligands of other Toll-like receptors to identify ligands which are specific for polypeptides of the invention. Preferably, the assays of the invention are carried out under conditions which would result in immunomodulatory NFkB mediated activity in the absence of the test substance, to identify inhibitors of Toll-like receptor mediated activity, or agents which inhibit ligand-induced Toll-like receptor activity.

25

21

Suitable test substances which can be tested in the above assays include combinatorial libraries, defined chemical entities, peptides and peptide mimetics, oligonucleotides and natural product libraries, such as display (e.g. phage display libraries) and antibody products. In a preferred embodiment, the test substance is a variant peptide of the invention. In particularly preferred embodiments, suitable test substances which may be candidate ligands for binding to and modulation of TLR9 or TLR9-A include synthetic or naturally occurring oligonucleotides or polynucleotides which bind directly to the Toll-like receptor or which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

10

The assay may be carried out using full length TLR9 to identify a variant peptide which interferes with TLR9 mediated activity, for example by inhibiting ligand binding.

15

Test substances may be used in an initial screen of, for example, 10 substances per reaction, and the substances of these batches which show inhibition or activation tested individually. Test substances may be used at a concentration of from 1nM to 1000µM, preferably from 1µM to 100µM, more preferably from 1µM to 10µM.

20

A protein-binding assay may be developed using a polypeptide of the invention, preferably one encoding the extracellular ligand-binding domain, to identify novel protein ligands of TLR8. Particular examples may be screening of a human cDNA expression library for protein ligands of TLR9 by yeast 2-hybrid protein interaction assay (e.g. as described in International Patent Application No. WO89/49284).

25

30

22

Another aspect of the present invention is the use of polynucleotides encoding the TLR9 polypeptides of the invention to identify mutations in TLR9 genes which may be implicated in human disorders or to identify cells in which TLR9 is expressed. Identification of such mutations may be used to assist in diagnosis of immune system, lung, kidney, heart or other disorders or susceptibility to such disorders and in assessing the physiology of such disorders. In particular the polynucleotides of the invention may assist in diagnosis of asthma and rheumatoid arthritis. For example, a SNP (single nucleotide polymorphism) has been identified in the genomic DNA encoding TLR9 (G/A nucleotide: The SNP Consortium database accession number TSC0164834). This single base pair change lies in the DNA region encoding the 23 N-terminal residues of TLR9, and this region is spliced out of the mRNA encoding TLR9-A. The nucleotide at this SNP position may affect the efficiency of mRNA splicing in the two different variants - a G at this position may possibly disrupt the splicing machinery and an A might lead to more efficient splicing. Additionally, the presence of a G as compared to an A in an unspliced mRNA would introduce a stop codon and result in different N terminal protein sequences upon translation of that mRNA, thus the two polymorphic variants of the *tlr9* gene may encode receptors which have differing expression levels and/or differing functional activity levels. The present invention provides a diagnostic tool for determining the polymorphic variant in an individual by detecting the DNA sequence at the SNP site. Such a tool may incorporate a nucleotide probe specific for one or other of the polymorphic variants, for example an oligonucleotide of from 5 to 50, preferably 5-20 nucleotides, complementary to a fragment of the nucleotide sequence of SEQ ID No. 1 which extends over the SNP site or a fragment complementary to that sequence with the exception of the single nucleotide change (G to A) at the SNP site. The present invention also provides a method of detecting a polymorphic variant in the *tlr9* gene by determining the sequence of nucleotides at and around the SNP site identified by the SNP consortium database accession number

23

TSC0164834, in particular by determining whether the nucleotide at that SNP site is a G or an A.

Another aspect of the present invention is the use of the compounds that have been identified by screening techniques referred to above in the treatment or prophylaxis of disorders which are responsive to regulation of TLR9 receptor activity. In addition, variant peptides of the invention which inhibit TLR9-mediated activity, for example which inhibit ligand binding or prevent hTLR9 immunomodulatory activity may be used in the treatment or prophylaxis of such disorders. Antibodies which recognise TLR9 may similarly be used in therapy.

In particular, such compounds may be used in the treatment of inflammatory, cardiovascular, systemic infection or autoimmune disease. The compounds may be used to treat bacterial, viral or fungal infections, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis.

In an alternative aspect, the invention provides agents which activate TLR9 mediated immunomodulation for use as an immunoadjuvant, or TLR9, and variants thereof, or polynucleotides or DNA encoding a polypeptide of the invention which may be administered for use as immunoadjuvants to enhance or alter the immune response in an individual to an antigen.

The compounds identified according to the screening methods outlined above may be formulated with standard pharmaceutically acceptable carriers and/or excipients as is routine in the pharmaceutical art, and as fully described in Remington's Pharmaceutical Sciences, Mack Publishing Company, Eastern

Pennsylvania 17th Ed. 1985, the disclosure of which is included herein of its entirety by way of reference.

5 The compounds may be administered by enteral or parenteral routes such as via oral, buccal, anal, pulmonary, intravenous, intra-arterial, intramuscular, intraperitoneal, topical or other appropriate administration routes. The dose of a compound to be administered may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the patient to be treated; the route of administration; and the required regimen. Again, a physician will be able to determine the required route of administration and dosage for any particular patient. A typical daily dose is from about 0.1 to 50 mg per kg of body weight, according to the activity of the compound, the age, weight and conditions of the subject to be treated, the type and severity of the disease and the frequency and route of administration.

15 Preferably, daily dosage levels are from 5 mg to 2 g.

Nucleic acid encoding an inhibitor of TLR9 activity may be administered to the mammal. In an alternative aspect of the invention, nucleic acid encoding TLR9 or a variant thereof may be administered to the animal. Such a variant shows immunomodulatory activity of TLR9 such as inducing cytokine production and expression of cell surface co-stimulatory receptors required for activation of T-cells. Nucleic acid, such as RNA or DNA, and preferably, DNA, is provided in the form of a vector, such as the polynucleotides described above, which may be expressed in the cells of the mammal.

25 Nucleic acid encoding the peptide may be administered to the animal by any available technique. For example, the nucleic acid may be introduced by injection, preferably intradermally, subcutaneously or intramuscularly. Alternatively, the nucleic acid may be delivered directly across the skin using a nucleic acid delivery device such as particle-mediated gene delivery. The

nucleic acid may be administered topically to the skin, or to the mucosal surfaces for example by intranasal, oral, intravaginal, intrarectal administration.

5 Uptake of nucleic acid constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents includes cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the nucleic acid to be administered can be altered. Typically the nucleic acid is administered in the range of 1pg to 1mg, preferably to 1pg to 10µg nucleic acid for particle mediated gene delivery and 10µg to 1mg for other routes.

Examples

15 Example 1 - Screening for compounds which exhibit protein modulating activity

Mammalian cells, such as HeK293, CHO and HeLa cells over-expressing TLR9 protein are generated for use in the assay. 96 and 384 well plate, high throughput screens (HTS) are employed. TLR9 induced cell activation can be monitored through activation of a signal transduction event (typically resulting in activation of NFκB or AP-1) or transcriptional activation of a reporter gene (typically regulated via NFκB or AP-1 responsive elements). TLR9 induced activation of a reporter gene such as alkaline phosphatase or luciferase can be easily assessed using a suitable colourimetric or fluorimetric assay to measuring production. Such assays may be used to identify receptor antagonists capable of blocking TLR9 activation, inhibitors capable of blocking TLR9 intracellular signal transduction or receptor agonists capable of activating TLR9. Secondary screening involves evaluation of TLR9 modulators in disease related

26

cells. Tertiary screens involve the study of modulators in rat and mouse models of disease relevant to the target.

27

SEQUENCE LISTING

5 <110> GLAXO GROUP LIMITED
<120> PROTEIN
<130> P78721 GCM SER
10 <140>
<141>
<160> 2
15 <170> PatentIn Ver. 2.1
<210> 1
<211> 3165
<212> DNA
<213> Homo sapiens
20 <220>
<221> CDS
<222> (1)..(3165)
25 <400> 1
atg ccc atg aag tgg agt ggg tgg agg tgg agc tgg ggg ccg gcc act 48
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Gly Pro Ala Thr
1 5 10 15
cac aca gcc ctc cca ccc cca cag ggt ttc tgc gcc agc gcc ctg cac 96
His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His
30 20 25 30
ccg ctg tct ctc ctg ctg cag gcc atc atg ctg gcc atg acc ctg gcc 144
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
ctg gtt acc ttg cct gcc ttc cta ccc tgc ggg ctg cag ccc cac gcc 192
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Gln Pro His Gly
50 55 60
ctg ctg aac tgc aac tgg ctg ttc ctg aag tct gtg ccc cac ttc tcc 240
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
40 65 70 75 80
atg gca gca ccc cgt gcc aat gtc acc agc ctt tcc ttg tcc tcc acc 288
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
85 90 95
gpc atc cac cac cat gat tct gac ttt gcc cac ctg ccc agc ctg 336
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
cgg cat ctc aac ctc aag tgg aac tgc ccg ccg gtt gcc ctg agc ccc 384
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
atg cgc ttc ccc tgc cac atg acc atc ggg ccc agc acc ttc ttg gct 432
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
130 135 140

Gln Thr Leu Arg Leu Gln Met Asn Phe 11e Asn Gln Ala Gln Leu Gly 420 425 430
 atc ttc agg gcc ttc cct ggc ctg tac ggc ctg ctg gac aac 1344
 11e Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn 435 440 445
 5 cgc atc agc gga gct tgg ggg ctg aca gcc acc atg ggg gga gca gat 1392
 Arg 11e Ser Gly Ala Ser Gly Thr Ala Thr Met Gly Ala Asp 450 455 460
 10 gga ggg ggg aag gtc tgg ctg cag cct ggg gac ctt gct ctg gcc cca 1440
 Gly Gly Gly Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro 465 470 475 480
 ggc gac act ccc agc tct gaa gac ttc agg ccc aac tgc agc acc ctc 1408
 Val Asp Thr Pro Ser Ser Gly Asp Phe Arg Pro Asn Cys Ser Thr Leu 485 490 495 495
 15 aac ttc acc ttg gac ctg tca cgg aac aac ctg ggc acc ggc cag cag 1536
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro 500 505 510 510
 gag atg ttt gcc cag ctg ctg cag ctg ggc tgc ctg ggc agc cag 1584
 Gln Met Phe Ala Gln Ser His Leu Gln Cys Leu Arg Leu Ser His 515 520 525 525
 20 aac tgc atc tgg ctg gca gtc atc ggc tcc cag ttc ctg ctg acc 1632
 Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr 530 535 540 540
 25 ggt ctg cag gtc cta gac ctg tcc cag aat agc ctg gac ctc tac cag 1680
 Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 545 550 555 560
 gag cac tca ttc agg ggg cta cca cga ctg gag gct ctg gac ctc agc 1728
 Gln His Ser Phe Thr Gln Leu Pro Arg Leu Gln Ala Leu Asp Leu Ser 565 570 575 575
 30 tac aac agc cag ccc ttt ggc atg cag ggc ctg ggc cag aac ttc agc 1776
 Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser 580 585 590 590
 ttc ggc gct cac ctg cgc acc ctg cgc cag ctc agc ctg ggc cag aac 1824
 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 595 600 605 605
 35 aac atc cag agc cca ggc tcc cag cag ctc tgc agt agc tgc ctg cgg 1872
 Asn 11e His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg 610 615 620 620
 ggc ctg gac ttc agc ggc aat gca ctg ggc cct atg tgg gcc gag gga 1920
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Gly Gly 625 630 635 640
 gac ctc tat ctg cag ttc ttc caa ggc ctg agc ggt ttc atc tgg ctg 1968
 Asp Leu Tyr Leu His Phe Gln Gly Leu Ser Gly Leu 11e Trp Leu 645 650 655 655
 45 gac ttg tcc cag aac cgc ctg cag acc ctc ctg ccc cag acc ctg cgc 2016
 Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg 660 665 670 670
 aac ctc ccc aag agc cta cag ggc ctg cgc ctc ctc ggc aat tac ctg 2064
 Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu 675 680 685 685
 50 ggc ttc ttt aag tgg tgg agc ctc cag ttc ctg ccc aac ctg gaa gtc 2112
 Ala Phe Phe Lys Trp Ser Leu His Phe Leu Pro Lys Leu Gln Val 690 695 700 700

gag ccc acc ctg gaa ggg cta aac cag agc tac aac aac atc atg act 480
 Val Pro Thr Leu Gln Gln Leu Ser Tyr Asn Asn 11e Met Thr 155 160 160
 146 ggc cct ggc ctg ccc aaa tcc cta ttc ctg tcc ctc agc cag acc 528
 5 Val Pro Ala Leu Pro Lys Ser Leu 11e Ser Leu Ser Leu Ser His Thr 165 170 175 175
 aac atc ctg atg cta gac tct ggc acc ctc ggc ctg cag ctc ctg 576
 Asn 11e Leu Met Leu Asp Ser Leu Ala Gln Leu His Ala Leu 180 185 190 190
 10 cgc ttc cta ttc atg gac ggc aac tot tat tac aag aac ccc tgc agg 624
 Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Lys Asn Pro Cys Arg 195 200 205 205
 cag gca ctg ggg ggc ggc ctg ggc ctc ctt ggc ctg ggc aac ctc 672
 Gln Ala Leu Gln Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu 210 215 220 220
 15 acc cag ctg tca ctc aag tac aac aac ctc act ggc ggc ccc cgc aac 720
 Thr His Leu Ser Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn 225 230 235 240
 20 ctg cct tcc agc ctg ggg tat ctg ctg ttc tac aac cgc atc gtc 768
 Leu Pro Ser Ser 245 250 255 255
 aac ctg ggc cct ggc gac ctg ggc aat ctg acc ggc ctg ggc ctc 816
 Lys Leu Ala Pro Gln Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 260 265 270 270
 25 ggt ggc ggc gga aat tgc cgc agc tgc gac cag ctc acc ccc cgc tgc 864
 Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys 275 280 285 285
 atg gag tgc cct ggc ttc ctc cag cta cat ccc gac acc ttc agc 912
 Met Gln Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser 290 295 300 300
 30 cag ctg agc ggc ctt gaa ggc ctg ggc agc agt tct ctc ttc 960
 His Leu Ser Arg Leu Gln Gly Leu Val Leu Lys Asp Ser Leu Ser 305 310 315 320
 35 tgg ctg aat ggc atg tgc ttc ggc ggc ggc aac ctc cga ggc ctg 1008
 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 325 330 335 335
 gac ctg agt ggc aac ttc ctc tcc aag tgc atc act aac acc agc gcc 1056
 Asp Leu Ser Gln Asn Phe Leu Tyr Lys Cys 11e Thr Lys Thr Lys Ala 340 345 350 350
 40 ttc cag ggc cta aca cag ctg cgc agc ctt aac ctg ttc ttc aat tac 1104
 Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 355 360 365 365
 caa aag agg ggc tcc ttt gcc ctc ctg tct ctg gcc cct tcc ttc ggg 1152
 Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly 370 375 380 380
 45 agc ctg gtc gcc ctg aag ggc ctg gac agc ggc atc ttc ttc cgc 1200
 Ser Leu Leu Ala Leu Lys Gln Leu Asp Met His Gly 11e Phe Phe Arg 385 390 395 400
 tca ctc atg agc acc cgc ctc cgc cgc ctg ggc ctg ccc atg ctc 1248
 50 Ser Leu Asp Gln Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 405 410 415 415
 cag act ctg cgc ctg agc atg aac ttc atc aac cag ggc cgc ctc ggc 1296

30

690 ctc cag gca gga aac cag cty aag gcc cty acc aat gcc acc cty 2160
Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Thr Asn Gly Ser Leu
705 710 715 720
5 ctc gcc gcc acc cag ctc cag gga cty gat gcc acc gcc acc ctc 2208
Pro Ala Gly Thr Arg Leu Arg Leu Asp Val Ser Cys Asn Ser Ile
725 730 735
9gc ttc gty gcc ccc gcc ttc ttc aag gcc aag gaa cty cga gaa 2256
Ser Phe Val Pro Gly Phe Ser Lys Ala Lys Glu Leu Arg Glu
740 745 750
ctc aac ctt acc gcc acc gcc ctc aag acc gty gcc ctc ttc ttc 2304
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
755 760 765
ggg ccc cty gga ggt gcc cty cca ata cta gat gta acc gcc acc cct 2352
15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro
770 775 780
ctc ctc tgc gcc tct ggg gcc gcc ttt atg gcc ttc cty cty gaa gty 2400
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu Val
785 790 795
cag gcc gcc gty ccc gcc cty ccc acc cag gga gty aag tct gcc acc cag 2448
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
805 810 815
ggc cag ctc cag gcc ctc acc ttt gcc cag gcc cty gcc ctc tgc 2496
25 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Leu Cys
820 825 830
ctg gat gaa gcc ctc ttc tgg gcc tct ttc gcc ctc ttc cty cty gcc 2544
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
835 840 845
gty gcc cty gcc cty gty cty cty cty cty cty cty cty cty cty 2592
30 Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp
850 855 860
ggc ctc tgg tac tgc ttc ctc ctc cty gcc tgg cty ctc ctc tgg cga 2640
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
865 870 875
35 ggg cga caa agt ggg cga gat gaa gcc gcc cty ccc tac gcc gcc ttc 2688
Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
885 890 895
gty gcc ttc gcc aaa acc cag gcc gcc gty gcc tgg gty tac aac 2736
40 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn
900 905 910
gag ctt cgg ggg cag cty gaa gaa tgc cgt ggg gcc tgg gca ctc cgc 2784
Glu Leu Arg Gly Glu Leu Glu Glu Cys Arg Gly Arg Trp Ala Leu Arg
915 920 925
ctg tgc cty gaa gaa gcc gcc tgg cty cct gcc aaa acc ctc ttc gaa 2832
45 Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
930 935 940
aac cty tgg gcc tgg gty tct gcc gcc acc aag acc cty ttc gty cty 2880
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
945 950 955
ggc cac acc gcc gcc gcc gcc ttc cty cty gcc gcc acc ttc cty cty 2928
60 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
965 970 975

31

ggc cag cag gcc cty cty gaa gcc gcc acc gcc gcc gty gty cty gty 2976
Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
980 985 990
atc cty gcc cct gcc gcc gcc gcc ctc ttc gcc tct gty cgg cty gcc cag 3024
5 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Arg Glu
995 1000 1005
cgc ctc tgc gcc cag agt gcc ctc ctc tgc ccc cag gcc ccc agt ggt 3072
Arg Leu Cys Arg Glu Ser Val Leu Leu Trp Pro His Glu Pro Ser Gly
1010 1015 1020
10 cag gcc gcc ttc tgg gcc cty gcc atg gcc acc agg gcc aac 3120
Gln Arg Ser Phe Trp Ala Glu Glu Gly Met Ala Leu Thr Arg Asp Asn
1025 1030 1035
cac gcc ttc tat aac cgg acc ttc tgc cag gga ccc acc gcc gaa 3165
15 His His Phe Tyr Asn Arg Asn Phe Cys Glu Gly Pro Thr Ala Glu
1045 1050 1055
<21> 2
<21> 1055
<22> PRT
<23> Homo sapiens
20
<40> 2
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr
1 5 10 15
25 His Thr Ala Leu Pro Pro Pro Glu Gly Phe Cys Arg Ser Ala Leu His
20 25 30
Pro Leu Ser Leu Leu Val Glu Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Glu Pro His Gly
50 55 60
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
65 70 75
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
80 85 90
35 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
130 135 140
40 Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
145 150 155
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
165 170 175
45 Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
180 185 190
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
195 200 205
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
210 215 220
50 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
225 230 235
240

Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val 250
 245
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 255
 260
 5 Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys 265
 270
 275
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser 280
 290
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser 300
 310
 10 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 315
 320
 325
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala 330
 340
 15 Phe Glu Gly Leu Thr Glu Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 345
 350
 355
 Glu Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly 360
 370
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg 375
 380
 20 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 385
 390
 400
 405
 Glu Thr Leu Arg Leu Glu Met Asn Phe Ile Asn Glu Ala Glu Leu Gly 410
 420
 25 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asn 425
 430
 435
 Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp 440
 445
 450
 Gly Gly Glu Lys Val Trp Leu Glu Pro Gly Asp Leu Ala Pro Ala Pro 455
 460
 30 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu 465
 470
 475
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Glu Pro 480
 485
 490
 500
 35 Glu Met Phe Ala Glu Leu Ser His Leu Glu Cys Leu Arg Leu Ser His 505
 510
 515
 Asn Cys Ile Ser Glu Ala Val Asn Gly Ser Glu Phe Leu Pro Leu Thr 520
 525
 530
 Gly Leu Glu Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 535
 540
 40 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser 545
 550
 555
 Tyr Asn Ser Glu Pro Phe Gly Met Glu Val Gly His Asn Phe Ser 560
 565
 570
 45 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 575
 580
 585
 590
 Asn Ile His Ser Glu Val Ser Glu Glu Leu Cys Ser Thr Ser Leu Arg 595
 600
 605
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly 610
 615
 50 Asp Leu Tyr Leu His Phe Phe Glu Glu Ser Gly Leu Ile Trp Leu 620
 625
 630
 635
 640
 645
 650
 655

Asp Leu Ser Glu Asn Arg Leu His Thr Leu Leu Pro Glu Thr Leu Arg 660
 665
 Asn Leu Pro Lys Ser Leu Glu Val Leu Arg Leu Arg Asp Asn Tyr Leu 670
 675
 5 Ala Phe Phe Lys Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 680
 685
 690
 Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Leu Thr Asn Gly Ser Leu 695
 700
 705
 Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile 710
 715
 720
 Ser Phe Val Ala Pro Gly Phe Ser Lys Ala Lys Glu Leu Arg Glu 725
 730
 735
 740
 Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe 745
 750
 755
 15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro 760
 765
 770
 Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu Val 775
 780
 785
 Glu Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro 790
 795
 800
 20 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Arg Leu Cys 805
 810
 815
 820
 Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala 825
 830
 835
 25 Val Ala Leu Gly Leu Gly Val Pro Met Leu His Leu Cys Gly Trp 840
 845
 850
 Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg 855
 860
 865
 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe 870
 875
 880
 30 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn 885
 890
 900
 Glu Leu Arg Gly Glu Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 905
 910
 915
 35 Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu 920
 925
 930
 Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu 935
 940
 945
 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu 950
 955
 40 Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val 960
 965
 970
 975
 980
 985
 990
 1000
 1005
 45 Arg Leu Cys Arg Glu Ser Val Leu Leu Trp Pro His Glu Pro Ser Gly 1010
 1015
 1020
 Glu Arg Ser Phe Thr Ala Glu Leu Gly Met Ala Leu Thr Arg Asp Asn 1025
 1030
 1035
 His His Phe Tyr Asn Arg Asn Phe Cys Glu Gly Pro Thr Ala Glu 1040
 1045
 1050
 50 1055

34

35

Seq. Id. no. 3			
atg ggt ttc tgc cgc agc gcc ctg cac		atg gag tgc cct cgt cac ttc ccc cag cta cat ccc ggt acc ttc agc	843
met gty phe cys arg ser ala leu his		met glu cys pro arg his phe pro gin leu his pro asp thr phe ser	
5	5		
ctg ctg tct ctg ctg ctg cag gcc atc atg ctg gcc atg acc ctg gcc		ctc cag ggc cta acc cag ctg cgc aag ctt aac ctg ttc ttc atc tac	1035
pro leu ser leu leu val gin ala ile met leu ala met thr leu ala		phe gin gly leu thr gin leu arg lys leu asn leu ser phe asn tyr	
10	10		
ctg ggt acc ttg cct gcc ttc cta ccc tgt gag ctg cag ccc cag gcc		330	335
leu gty thr leu pro ala phe leu pro cys gln leu leu pro his gty		335	340
30	35		
ctg gty aac tgc aac tgg ctg ttc ctg aag tct gty ccc cag ttc ttc		gac ctg agt gag aac ttc ctc tac aaa tgc atc act aaa acc aag gcc	987
leu val asn cys asn trp leu phe leu lys ser val pro his phe ser		asp leu ser glu asn phe leu tyr lys cys ile thr lys thr lys ala	
45	50		
atg gca gca ccc cgt gcc aat gtc acc gcc cct tcc ttg tcc tcc aac		315	320
met ala ala pro arg gty asn val thr ser leu ser leu ser ser asn		320	325
60	65		
cgc atc cac cac ctc cat ggt tct gac ttt gcc cag ctg ccc agc ctg		325	330
arg ile his his leu his asp ser asp phe ala his leu pro ser leu		330	335
75	80		
cgg cat ctc aac ctg aag tgg aac tgc cgg cgt gcc ctg ctc agc ccc		335	340
arg his leu asn leu lys trp asn cys pro pro val gty leu ser pro		340	345
90	95		
atg aac ttc ccc tgc cac atg acc atc gag ccc agc acc ttc ttg ggt		345	350
met his phe pro cys his met thr ile glu pro ser thr phe leu ala		350	355
105	110		
gty ccc acc ctg gaa gag cta aac ctg agc tac aac aac atc atg act		355	360
val pro thr leu glu leu leu leu ser tyr asn asn ile met thr		360	365
115	120		
gty cct ggg ctg ccc aaa tcc ctc ata tcc ctg tcc ctg agc cat acc		365	370
val pro ala leu pro lys ser leu ile ser leu ser leu ser his thr		370	375
125	130		
aac atc ctg atg cta gcc tct gcc gcc ctc gcc gcc ctg cat gcc ctg		375	380
asn ile leu met leu asp ser ala ser leu ala ile leu his ala leu		380	385
135	140		
cgc ttc cta ttc atg gac gcc aac tgt tat tac aag aac ccc tgc agg		385	390
arg phe leu phe met asp gty asn cys tyr tyr lys asn pro cys arg		390	395
145	150		
cag gca ctg gag gty gcc cgg ggt gcc ctc ctt gcc ctg gcc aac ctc		395	400
gin ala leu glu val ala pro gty ala leu leu gty leu gty asn leu		400	405
155	160		
acc cac ctg tca ctc aag tac aac aac ctc act gty gty ccc gcc aac		405	410
thr his leu ser leu lys tyr asn asn leu thr val val pro arg asn		410	415
165	170		
ctg cct acc ctg gag tat ctg ctg ttg ttc tac aac gcc atc gtc		415	420
leu pro ser ser leu glu tyr leu leu leu ser tyr asn arg ile val		420	425
175	180		
aaa ctg ggg cct gag gcc ctg gcc aat ctg acc gcc ctg ggt gcc		425	430
lys leu ala pro glu asp leu ala asn leu thr ala leu arg val leu		430	435
185	190		
ggt gty ggc gga aat tgc cgc gcc tgc gcc cac gct ccc aac ccc tgc		435	440
asp val gty gty asn cys arg arg cys asp his ala pro asn pro cys		440	445
195	200		
205	210		
215	220		
225	230		
235	240		
245	250		
255	260		
265	270		
275	280		
285	290		
295	300		
305	310		
315	320		
325	330		
335	340		
345	350		
355	360		
365	370		
375	380		
385	390		
395	400		
405	410		
415	420		
425	430		
435	440		
445	450		
455	460		
465	470		
475	480		
485	490		
495	500		
505	510		
515	520		
525	530		
535	540		
545	550		
555	560		
565	570		
575	580		
585	590		
595	600		
605	610		
615	620		
625	630		
635	640		
645	650		
655	660		
665	670		
675	680		
685	690		
695	700		
705	710		
715	720		
725	730		
735	740		
745	750		
755	760		
765	770		
775	780		
785	790		
795	800		
805	810		
815	820		
825	830		
835	840		
845	850		
855	860		
865	870		
875	880		
885	890		
895	900		
905	910		
915	920		
925	930		
935	940		
945	950		
955	960		
965	970		
975	980		
985	990		
995	1000		

37

815 820 825
gtg ctc ctc ggc ctc ggc ctc atg ctc ctc ctc ctc ctc ggc tgg 2523
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp
5
gac ctc tgg tac tgc ttc ctc ctc ctc tgc tgc tgc ctc ctc ctc tgg cgg 2571
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850 855
ggg cgg caa agt ggg cga ggt ggc ctc ctc ctc ctc ctc ggc ttc 2619
Gly Arg Gln Ser Gly Arg Asp Ala Leu Pro Tyr Asp Ala Phe
860 865 870
gtg gtc ttc gac aaa acg cag agc gca gtc gca gtc tgg gtc tac aac 2667
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn
875 880 885
gag ctt cgg ggg cag ctc ggc ggc tgc cgt ggg cgc tgg gca ctc cgc 2715
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg
890 895 900
ctg tgc ctc ggc gaa cgc cgc tgg ctc ctc ggc aaa acc ctc ttt ggc 2763
Leu Cys Leu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
910 915
aac ctc tgg ggc tgc gtc tat ggc agc agc agc ctc ttt gtc ctc 2811
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
925 930 935
gcc cac acg gac cgg gtc gtc ctc ttc ctc ctc ggc agc ttc ctc ctc 2859
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
940 945 950
gcc cag cag cgc ctc gtc ggc agc agc agc gtc gtc gtc ctc gtc 2907
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val
955 960 965
atc ctc ggc ctc ggc cgc ctc ctc ctc tat ggc cgc ctc ctc ctc 2955
Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln
970 975 980
cgc ctc tgc cgc cgc agt gtc ctc ctc tgg ctc ctc cgc cgc cgc agt ggt 3003
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
990 1000
cgc cgc agc ttc tgg ggc cag ctc ggc agc ctc ctc ctc agc agc aac 3051
Gln Arg Ser Phe Trp Ala Gln Leu Glu Gly Met Ala Leu Thr Arg Asp Asn
1005 1010 1015
cac cac ttc tat aac cgg aac ttc tgc cag gga cgc agc ggc gaa 3096
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
1020 1025 1030

38

540 545 550
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
555 560 565
tac aac agc cgc ttt ggc atg cag ggc ggc cgc cgc aac ttc agc 1707
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
570 575 580
ttc ggc gtc ctc ctc agc cgc cgc ctc agc ctc ggc cgc cgc aac 1755
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
585 590 595
aac atc cgc cgc cgc cgc cgc cgc cgc cgc cgc cgc cgc cgc cgc 1803
Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
600 605 610
gcc ctc gtc ttc ggc agc aac gca ctc gtc ctc atg tgg ggc ggc gga 1851
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
615 620 625
gac ctc tat ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc 1899
Asp Leu Tyr Leu His Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe
630 635 640
gac ttc ttc cag aac agc ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc 1947
Asp Leu Ser Gln Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
645 650 655
aac ctc ctc ang agc ctc cag gtc ctc ctc ctc ctc ggc aac tac ctc 1995
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
660 665 670
gcc ttc ttc aag tgg tgg agc ctc ctc ttc ctc ctc aaa ctc gaa gtc 2043
Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val
675 680 685
ctc gac ctc gca gga aac cag ctc agc ctc agc ctc aac agc agc ctc 2091
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
690 695 700
ctc gtc agc ctc cgc cgc cgc agc ctc gtc agc tgc agc agc atc 2139
Pro Ala Gly Thr Arg Leu Arg Leu Asp Val Ser Cys Asn Ser Ile
705 710 715
agc ttc gtc ggc ctc ctc ttc ttc aag ggc aag ggc ctc cgc ggc 2187
Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu
720 725 730
ctc aac ctt agc cgc aac ggc ctc aag aca gtc ggc ctc ttc tgg ttc 2235
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
735 740 745
ggg ctc ctc ggc agt ggc ctc ctc ata ctc gta gta agc agc aac ctc 2283
Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
750 755 760
ctg ctc tgc gtc tgc tgc ggc ggc ttc atg gac ttc ctc ctc ggc gtc 2331
Leu His Cys Ala Cys Gly Ala Phe Met Asp Phe Leu Leu Glu Val
765 770 775
cag gtc gtc gtc ctc gtc ctc ctc ctc ctc ctc ctc ctc ctc ctc ctc 2379
Gln Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
780 785 790
gac cag ctc ctc ggc ctc agc atc ttc gca cag gac ctc ctc tgc 2427
Gly Gln Leu Gln Gly Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
795 800 805
ctg gtc ggc ctc ctc tgc gtc tgc ttc gtc ctc ctc ctc ctc ctc gtc 2475
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala

38

45 Met Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn 55
60 Arg 116 His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu 70
75 Arg His Leu Asn Leu Lys Trp Asn Cys Pro Val Gly Leu Ser Pro 85
90 Met His Phe Pro Cys His Met Thr 116 Glu Pro Ser Thr Phe Leu Ala 105
110 Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn 116 Met Thr 120
125 Val Pro Ala Leu Pro Lys Ser Leu 116 Ser Leu Ser Leu Ser His Thr 135
140 Asn 116 Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu. 150
155 Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg 165
170 Glu Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu 180
190 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn 195
202 Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg 116 Val 215
220 Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 230
235 Asp Val Gly Gly Asn Cys Arg Cys Asp His Ala Pro Asn Pro Cys 245
250 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser 260
270 His Leu Ser Arg Leu Glu Gly Leu Val Lys Asp Ser Leu Ser 280
285 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 295
300 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys 116 Thr Lys Thr Lys Ala 310
315 Phe Glu Gly Leu Thr Glu Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 325
330 Glu Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly 340
350 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly 116 Phe Phe Arg 360
365 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 375
380 Glu Thr Leu Arg Leu Glu Met Asn Phe 116 Asn Glu Ala Glu Leu Gly 390
395 116 Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn 405
410 Arg 116 Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp 420
430 Gly Gly Glu Lys Val Trp Leu Glu Pro Gly Asp Leu Ala Pro Ala Pro 435
445 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu 455

39

450 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Glu Pro 470
475 Glu Met Phe Ala Glu Leu Ser His Leu Glu Cys Leu Arg Leu Ser His 485
490 Asn Cys 116 Ser Glu Ala Val Asn Gly Ser Glu Phe Leu Pro Leu Thr 495
500 Gly Leu Glu Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 510
515 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser 520
525 Tyr Asn Ser Glu Pro Phe Gly Met Glu Gly Val Gly His Asn Phe Ser 530
540 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 545
555 Asn 116 His Ser Glu Val Ser Glu Glu Leu Cys Ser Thr Ser Leu Arg 560
570 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly 575
585 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu 116 Trp Leu 590
605 Asp Leu Ser Glu Asn Arg Leu His Thr Leu Leu Pro Glu Thr Leu Arg 610
620 Asn Leu Pro Lys Ser Leu Glu Val Leu Arg Leu Arg Asp Asn Tyr Leu 625
635 Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 640
655 Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Leu Thr Asn Gly Ser Leu 655
670 Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser 116 680
695 Ser Phe Val Ala Pro Gly Phe Ser Lys Ala Lys Glu Leu Arg Glu 700
715 Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe 710
720 Gly Pro Leu Ala Ser Ala Leu Glu 116 Leu Asp Val Ser Ala Asn Pro 725
735 Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Val 730
745 Glu Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro 740
755 Gly Glu Leu Glu Gly Leu Ser 116 Phe Ala Glu Asp Leu Arg Cys 755
765 Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala 760
775 Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp 770
780 Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg 785
795 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe 800
810 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn 815
820 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995

40

41

CLAIMS

- 875
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 885
890
895
900
905
Leu Cys Leu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
910
915
920
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
925
930
935
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
940
945
950
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
955
960
965
Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln
970
975
980
985
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
990
995
1000
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
1005
1010
1015
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
1020
1025
1030
- 5 1. An isolated Toll-like-receptor polypeptide consisting essentially of
(i) the amino acid sequence SEQ ID NO: 2;
(ii) a variant thereof which has immunomodulatory activity; or
(iii) a fragment of (i) or (ii) which has immunomodulatory activity.
- 10 2. A polypeptide according to claim 1 wherein the variant (ii) has at least 70% identity to the amino acid sequence of SEQ ID NO: 2.
- 15 3. A polypeptide according to claim 2 which has at least 70% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 70% of the full-length sequence provided by SEQ ID No.1 and exhibits toll-like receptor functionality.
- 20 4. A polypeptide according to claim 1 or claim 2 wherein the variant (ii) has at least 95% identity to the amino acid sequence of SEQ ID NO: 2.
- 25 5. A polypeptide according to claim 2 which has at least 95% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 60 contiguous amino acids and exhibits toll-like receptor functionality.
- 30 6. A polypeptide according to claim 1 wherein the fragment (iii) is a peptide of up to 150 amino acids in length and exhibits toll-like receptor functionality.
7. A polynucleotide encoding a polypeptide according to any one of claims 1-3.
8. A polynucleotide encoding a Toll-like receptor polypeptide which has immunomodulatory activity, which polynucleotide consists essentially of:
(a) the nucleic acid sequence of SEQ ID NO: 1;
(b) a sequence complementary thereto;

42

- (c) a sequence which hybridises under stringent conditions to a sequence as defined in (a) or (b);
(d) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a), (b) or (c); or
(e) a sequence having at least 60% identity to a sequence as defined in (a), (b), (c) or (d).

9. A polynucleotide according to claim 7 or claim 8 which is mRNA.

10. A polynucleotide according to claim 7 or claim 8 which is DNA.

11. A polynucleotide according to claim 7 or claim 8 which is cDNA.

12. An isolated toll-like receptor polypeptide which is obtainable by expression *in vivo* or *in vitro* of a polynucleotide according to claim 7 or claim 8.

13. A polypeptide according to claim 12 which has the structural features conserved amongst toll-like receptors.

14. An expression vector comprising a polynucleotide sequence according to any one of claims 7 to 11, which is capable of expressing a polypeptide according to any one of claims 1 to 3 or claim 12.

15. An expression vector according to claim 14 which is a plasmid, phage or viral vector.

16. A host cell comprising an expression vector according to claim 14 or claim 15.

43

17. A polyclonal or monoclonal antibody, or a chimera or fragment thereof, which is specific for a polypeptide according to any one of claims 1 to 3.

18. A method for identification of a compound that modulates Toll-like receptor activity, which method comprises contacting a polypeptide according to any one of claims 1 to 3 or claim 12 with a test substance and monitoring for immunomodulatory activity.

19. A compound which modulates Toll-like receptor activity and which is identifiable by a method according to claim 18.

20. A compound according to claim 19 which is a peptide or polypeptide.

21. A compound according to claim 19 which is an oligonucleotide or polynucleotide.

22. A method of treating a subject having an inflammatory or cardiovascular disorder, systemic infection or autoimmune disease that is responsive to Toll-like receptor modulation, which method comprises administering to said subject an effective amount of a compound according to any one of claims 19 to 21 or an antibody according to claim 17.

23. A method according to claim 22 wherein the disorder is a viral, fungal or bacterial infection, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome associated with systemic infection involving gram positive or gram negative bacteria.

44

24. A polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, for use in the treatment of an immune or inflammatory disorder.

5

25. A polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 for use as an adjuvant.

26. The use of a compound according to claim 19 in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

27. The use of a polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

28. The use of a polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 in the manufacture of an adjuvant or vaccine formulation.

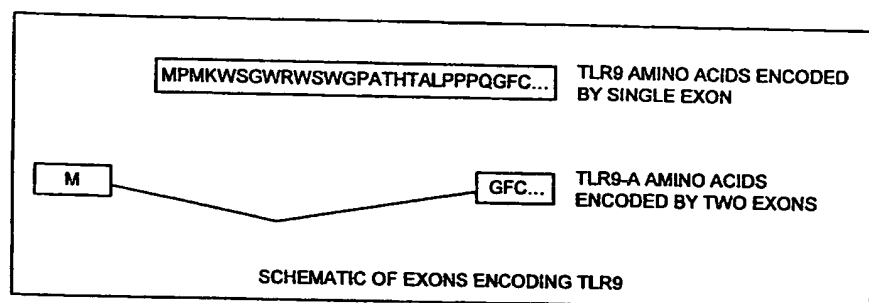
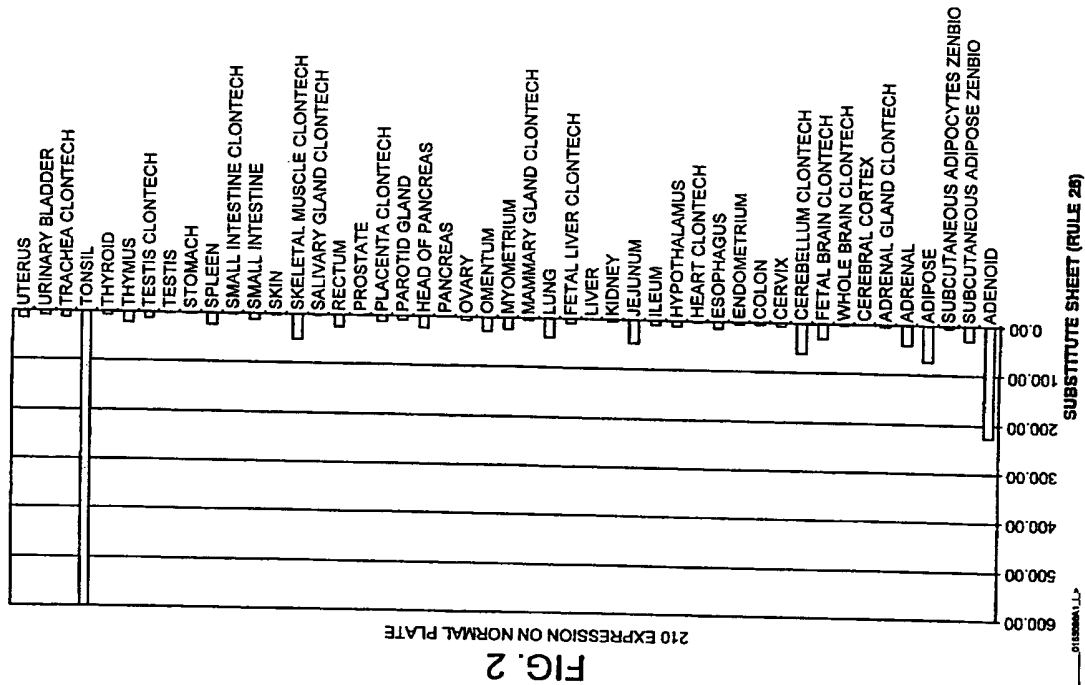
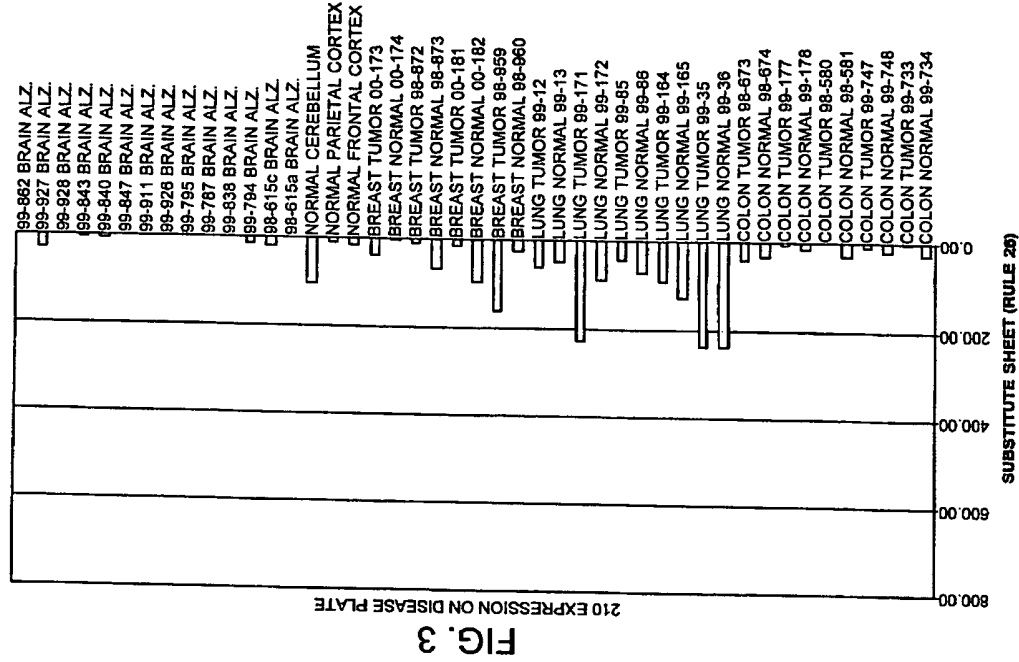


FIG. 1

2 / 5



3 / 5



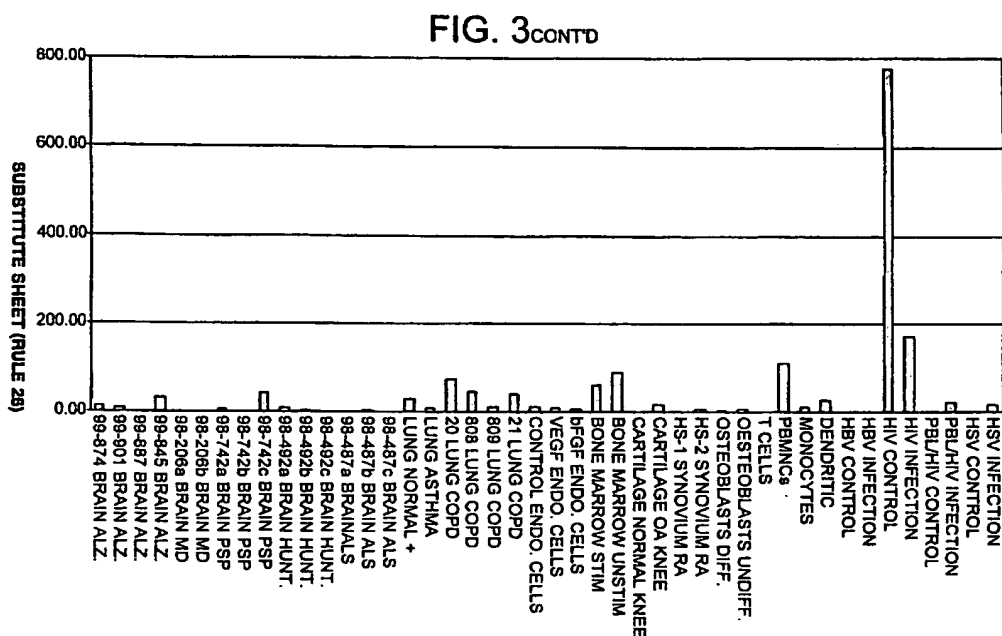
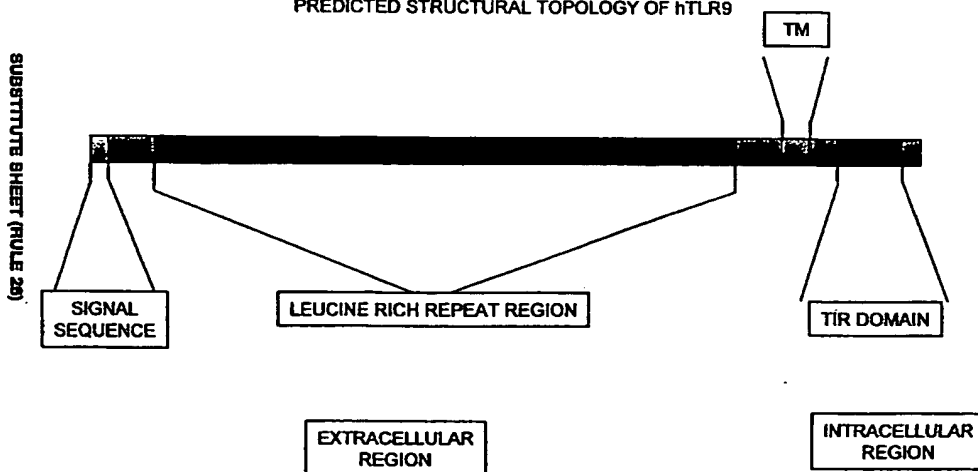


FIG. 4
PREDICTED STRUCTURAL TOPOLOGY OF hTLR9



INTERNATIONAL SEARCH REPORT

Intern. Appl. No. PCT/88 01/00299	
CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/63 C12N5/10 G01N33/68 C07K14/705 C07K16/28 A61K31/7088 A61K38/17 A61K48/00	
According to International Patent Classification (IPC) or to both national classification and IPC	
A. FIELDS SEARCHED Minimum documentation searched (classification symbols followed by classification symbols) IPC 7 C12N GOIN C07K A61K	
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched	
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EMBL, STRAND, EPO-Internal, NPI Data, PAJ, BIOSIS, MEDLINE, CHEM ABS Data	
C. DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with indication, where appropriate, of the relevant passages
X	WO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12) 99.8% identity in 1008 bp overlap between SEQ ID NO 33 of WO9850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of WO9850547 and SEQ ID NO 2 page 56, line 23 - page 60, line 14; claims 1-16; examples 1-11 -/-
Further documents are listed in the continuation of box C.	
* Special categories of cited documents: "A" document(s) defining the patent state of the art which is not considered to be of particular relevance "B" earlier document(s) published on or after the international filing date "C" document(s) which may have priority claims(s) or which may be considered to be of particular relevance to the invention "D" document(s) which may be considered to be of particular relevance to the invention "E" document(s) relating to an oral disclosure, use, exhibition or other means "F" document(s) published prior to the international filing date but later than the priority date claimed "G" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "H" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "I" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "J" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "K" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "L" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "M" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "N" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "O" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "P" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Q" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "R" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "S" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "T" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "U" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "V" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "W" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "X" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Y" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Z" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention	
Date of the actual completion of the international search 28 June 2001	
Name and mailing address of the ISA European Patent Office, P.O. Box 6516 Palamirien 2 NL - 2500 LV The Hague Tel. (+31-70) 340-3011 Fax (+31-70) 340-3015	
Authorized officer Devijver, K	

INTERNATIONAL SEARCH REPORT

Intern. Appl. No. PCT/88 01/00299	
C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with indication, where appropriate, of the relevant passages
X	DATABASE EMBL 'Online! accession: AC006252 29 December 1998 (1998-12-29) HUIZNY D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RFL15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407 100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)
P, X	HEMMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-745, XP002168474 ISSN: 0028-0836 cited in the application the whole document -8 DATABASE EMBL 'Online! accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for Toll-like receptor 9, complete cds." XP002170408 99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2
Further documents are listed in the continuation of box C.	
* Special categories of cited documents: "A" document(s) defining the patent state of the art which is not considered to be of particular relevance "B" earlier document(s) published on or after the international filing date "C" document(s) which may have priority claims(s) or which may be considered to be of particular relevance to the invention "D" document(s) which may be considered to be of particular relevance to the invention "E" document(s) relating to an oral disclosure, use, exhibition or other means "F" document(s) published prior to the international filing date but later than the priority date claimed "G" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "H" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "I" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "J" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "K" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "L" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "M" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "N" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "O" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "P" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Q" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "R" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "S" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "T" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "U" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "V" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "W" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "X" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Y" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention "Z" document(s) published after the international filing date but not in conflict with the application but which may be considered to be of particular relevance to the invention	
Date of the actual completion of the international search 11/07/2001	
Name and mailing address of the ISA European Patent Office, P.O. Box 6516 Palamirien 2 NL - 2500 LV The Hague Tel. (+31-70) 340-3011 Fax (+31-70) 340-3015	
Authorized officer Devijver, K	

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9850547	A	12-11-1998	AU 7175498 A BR 9808747 A CN 1263555 T EP 0980429 A NO 995458 A PL 336635 A SK 146589 A HU 0001462 A	27-11-1998 11-07-2000 16-08-2000 23-02-2000 08-11-1999 03-07-2000 11-07-2000 28-07-2000

Intern. Appl. No. PCT/88 01/00299

International Application No. PCT/88 01/00299

FURTHER INFORMATION CONTINUED FROM PCT/88/ 210

Continuation of Box I.2

Claims Nos.: 19-21,26; in part: 22,23,25,28

Claim 19 refers to a compound which modulates Tol-like receptor activity without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claim is ambiguous and vague, and its subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).

No meaningful search can be carried out for such a purely speculative claim whose wording is, in fact, a mere recitation of the result to be achieved. The above comment also applies for claims 20,21 and 26; and in part for claims 22,23,25 and 28.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

Form PCT/ISA/210 (patent family annex) (July 1997)

ENCLOSURE: 4PC 018288A 1,1

ENCLOSURE: 4PC 018288A 1,1

THIS PAGE BLANK (USPTO)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 August 2001 (02.08.2001)

(10) International Publication Number
WO 01/55386 A1

PCT

(51) International Patent Classification: C12N 15/02, 15/63, 510, G01N 33/68, C07K 14/703, 16/28, A61K 31/708, 39/17, 48/00

(52) International Application Number: PCT/GB01/00299

(53) International Filing Date: 25 January 2001 (25.01.2001)

(54) Filing Language: English

(55) Publication Language: English

(56) Priority Date: 25 January 2000 (25.01.2000) GB

(71) Applicant (for all designated States except US): GLAXO-SMITHKLINE (GB/GB); Corporate Intellectual Property, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

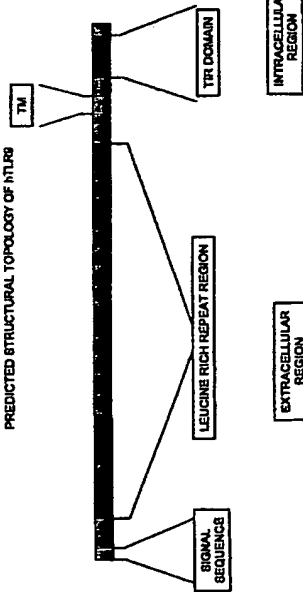
(72) Inventors and
(73) Investors/Applicants (for US only): LEWIS, Alan, Peter (GB/GB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB); RAY, Keith, Paul (GB/GB); Glaxo Wellcome plc, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY (GB).

Published:
— with international search report
— before the expiration of the time limits for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) TITLE: TOLL-LIKE RECEPTOR

PREDICTED STRUCTURAL TOPOLOGY OF hTLR9



(57) Abstract: An isolated Toll-like receptor polypeptide comprises the amino acid sequence of SEQ ID NO: 2, a variant or a fragment thereof which has immunomodulatory activity. Polynucleotides encoding such a Toll-like receptor are also described. A method for identification of a substance that modulates Toll-like receptor activity comprises contacting a polypeptide of the invention with a test substance and monitoring for immunomodulatory activity.

WO 01/55386

1

Toll-like Receptor

Field of the invention

The present invention relates to a novel Toll-like receptor or a variant thereof. A variant may demonstrate Toll-like receptor activity such as activation of NF- κ B, or may inhibit Toll-like receptor activity.

Background of the invention

A family of human Toll-like receptors has been described in the literature. These receptors are termed Toll-like receptors in view of common structural features shared with a Drosophila Toll (dToll) receptor molecule which is involved in embryonic development. Toll and Toll-like receptors are type I transmembrane proteins, with extracellular leucine rich repeat motifs and an intracellular signalling domain homologous to that of members of the interleukin 1 receptor superfamily.

Drosophila Toll also plays an important role in the adult fly and is involved in immune surveillance mechanisms required for recognition of bacterial and fungal pathogens and regulation of specific innate immune defence gene expression. Activation of dToll receptors in response to infection by specific micro-organisms is thought to require the production of a protein ligand called Spaetzle. The human Toll-like receptors (hTLRs) are also thought to participate in mechanisms of innate immunity and inflammation acting as pattern recognition receptors (PRRs) for bacteria and other micro-organisms. hTLRs are expressed on antigen presenting cells including monocytes and dendritic cells and like dToll play roles in innate immunity. TLRs can elicit pro-inflammatory cytokine production and induce expression of cell surface co-stimulatory receptors required for activation of T-cells. Some hTLRs may help to co-ordinate interactions between cells of the innate and acquired immune systems to orchestrate an integrated immune response to infection.

2

The specific functions of two mammalian TLRs, TLR2 and TLR4, have recently been identified. TLR2 and TLR4 are involved in mediating host responses to gram positive and gram negative bacteria through recognition of specific bacterial wall components. It has also recently been shown that TLR4 mediates responses to certain viral proteins including respiratory syncytial virus (RSV) (Nature Immunology 1: 398 2000).

Additionally TLRs may form heterodimeric functional complexes. Components of the intracellular signal transduction pathways of some hTLRs appear to be shared with Interleukin-1 (IL-1) receptor transduction pathways. Stimulation of TLR2 and TLR4 leads to activation of NF κ B via an adaptor protein MyD88 and recruitment of the IL-1 receptor associated kinases (IRAKs).

Summary of the invention

A novel Toll-like receptor is now provided which is a screening target for the identification and development of novel pharmaceutical agents which modulate the activity of the receptor and in particular have immunomodulatory activity. These agents may be used in the treatment and/or prophylaxis of inflammatory diseases, cardiovascular diseases, systemic infections and autoimmune diseases, such as asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, viral, fungal and bacterial infections, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis. These agents may also be used as immunoadjuvants to enhance or alter the immune response in vaccine therapy.

Accordingly, the present invention provides an isolated Toll-like-receptor polypeptide which comprises:

3

- (i) the amino acid sequence of SEQ ID NO: 2;
- (ii) a variant of (i) which has immunomodulatory activity; or
- (iii) a fragment of (i) or (ii) which retains immunomodulatory activity.

Preferably, a variant has at least 80% identity to the amino acid sequence of SEQ ID NO: 2, more preferably at least 95% identity therewith, for example 97% identity therewith.

The invention also provides a polynucleotide encoding a polypeptide of the invention. Such a polynucleotide may be a polynucleotide which encodes a Toll-like receptor polypeptide which has immunomodulatory activity. The polynucleotides of the invention may be DNA or RNA, for example mRNA. A polynucleotide according to the invention comprises:

- (a) the nucleic acid sequence of SEQ ID NO: 1 and/or a sequence complementary thereto;
- (b) a sequence which hybridises under stringent conditions to a sequence as defined in (a);
- (c) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a) or (b); or
- (d) a sequence having at least 60% percent identity to a sequence as defined in (a), (b) or (c).

The present invention also provides a polypeptide expressed from a polynucleotide according to (a), (b), (c) or (d) above, in particular a polypeptide comprising a toll-like receptor according to the invention, encoded by the mRNA derived from a DNA sequence according to (a) or (b) above, thus the invention provides an isolated toll-like receptor polypeptide which is obtainable by expression *in vitro* or *in vivo* of a DNA molecule comprising the sequence of nucleotides as shown in SEQ ID NO:1.

The polypeptides of SEQ ID NO 2 and SEQ ID NO 4, which are different isoforms expressed from the multiple exon *tlr9* gene, are herein referred to as TLR9 and TLR9-A, respectively. TLR9-A is encoded by the nucleotide sequence of SEQ ID No. 3, which is encoded within SEQ ID No.1 except for the initiating methionine, that is encoded by a second exon as illustrated in figure 1. (see Hemmi et al. Nature 408, 740-745 2000; Du et al, European Cytokine Network, 11: 362-371, 2000; Chuang and Ulevitch, European Cytokine Network, 11: 372-378 for isolation of the cDNAs, and corresponding sequence database accessions EMBL:AB045180, EMBL:AF259262, EMBL:AF245704).

10

In further aspects of the invention we provide:

- an expression vector capable of expressing a polypeptide of the invention comprising a polynucleotide as defined above.
- a host cell comprising an expression vector of the invention.
- an antibody specific for a polypeptide of the invention.
- a method for identification of a compound that modulates Toll-like receptor activity, comprising contacting a polypeptide of the invention with a test compound and monitoring for immunomodulatory activity.

15

20

Compounds which are identifiable in accordance with this method may be used in the treatment of a subject having a disorder that is responsive to Toll-like receptor modulation such as an inflammatory or cardiovascular disorder or systemic infection or autoimmune disease, including asthma, chronic obstructive pulmonary disease (COPD), emphysema, inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, viral, fungal and bacterial infections, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes and Multiple Sclerosis. In particular, compounds which are identifiable in accordance with this method may be used in the treatment of a subject having allergic asthma or

30

rhinitis. Further, such compounds may have immunomodulatory activity and be of use in the treatment of, or as adjuvants in vaccination against, bacterial or viral infections or as components of anti-cancer vaccines.

5

Compounds identifiable in accordance with this method include, in particular, synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which bind directly to the Toll-like receptor of the invention, and synthetic or naturally occurring oligopeptides or polypeptides, oligonucleotides or polynucleotides which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

10

In an alternative aspect of the invention, a polypeptide comprises a fragment or variant of SEQ ID NO 2 which is capable of inhibiting the activity of TLR9 or TLR9-A, for use in the treatment of an immune or inflammatory disorder.

15

In a further aspect of the invention, a polypeptide or polynucleotide in accordance with the invention or a compound identifiable in accordance with the invention is provided for use as an adjuvant or as an immunotherapeutic agent, for example in a vaccine.

20

Brief Description of the Sequences

SEQ ID NO: 1 is the amino acid sequence of human protein TLR9 and its encoding DNA;

25

SEQ ID NO: 2 is the amino acid sequence alone of TLR9;

SEQ ID NO: 3 is the amino acid sequence of human protein TLR9-A and its encoding cDNA (EMBL:AF259262);

SEQ ID NO: 4 is the amino acid sequence alone of TLR9-A (Hemmi et al.).

30

Brief Description of the Drawings

Fig. 1 is a diagrammatic illustration showing the exon arrangement encoding TLR9 and TLR9-A;

Fig. 2 shows tissue distribution data for TLR9 (using a human tissue cDNA plate). The profile shows predominant expression in tonsil and adenoid tissues with lower levels of expression detected in adipose, adrenal, foetal brain, cerebellum, jejunum, lung, myometrium, omentum, head of pancreas, rectum, skeletal muscle, spleen and thymus tissues;

Fig. 3 shows tissue distribution data for TLR9 (using a human disease cDNA plate). The profile shows predominant expression in lung tissue, bone marrow and PBMC with lower levels of expression detected in some colon, breast and brain/cerebellum samples;

Fig. 4 illustrates, in diagrammatic form, the predicted structural topology of human TLR9 - "TM" is the transmembrane portion, "TIR" is the cytosolic region conserved among Interleukin and toll-like receptors known as the Toll Interleukin Receptor domain.

Detailed Description of the Invention

A single open reading frame was identified in genomic DNA (SEQ ID No.1), which encodes a protein of 1055 residues, predicted molecular weight 118,515 (PeptideSort - GCG Software) and the amino acid sequence shown in SEQ ID No.2. This sequence included the TIR domain common to Toll-like receptors and members of the interleukin-1 receptor family e.g. IL1RI, and the N-terminal sequence contains structural features as shown in figure 4. These features include, in order from the N- to the C-terminus, a predicted signal sequence with a potential cleavage site between residues 48 and 49 or 50 and 51 (SPScan in GCG; SignalP), a leucine-rich repeat motif domain, a potential transmembrane region and the Toll/IL-1R homologous region (TIR; Pfam Database). These motifs confirmed that the protein was likely to be expressed as a single transmembrane receptor-like molecule belonging to the TLR rather than the IL-1R

family and therefore, it was designated TLR9 based on existing published and in-house nomenclature.

Throughout the present specification and the accompanying claims the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

The present invention relates to a human Toll-like receptor, referred to herein as TLR9, and variants or fragments thereof. Sequence information for TLR9 is provided in SEQ ID NO: 1 (nucleotide and amino acid) and in SEQ ID NO: 2 (amino acid only). A polypeptide of the invention consists essentially of the amino acid sequence of SEQ ID NO: 2 or of a functional variant of that sequence. One important variant of TLR9 is TLR9-A, sequence information for which is provided in SEQ ID NO: 3 (nucleotide and amino acid) and in SEQ ID NO: 4 (amino acid only).

The polypeptides are provided in isolated form. The term "isolated" is intended to convey that the polypeptide is not in its native state, insofar as it has been purified at least to some extent or has been synthetically produced, for example by recombinant methods. The polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. The term "isolated" therefore includes the possibility of the polypeptide being in combination with other biological or non-biological material, such as cells, suspensions of cells or cell fragments, proteins, peptides, expression vectors, organic or inorganic solvents, or other materials where appropriate, but excludes the situation where the polypeptide is in a state as found in nature.

A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the invention. Routine methods, can be employed to purify and/or synthesise the proteins according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2nd Edition, CSH Laboratory Press (1989), the disclosure of which is included herein in its entirety by way of reference.

10

The term "variants" refers to polypeptides which have the same essential character or basic biological functionality as TLR9. The essential character of TLR9 can be defined as that of a Toll-like receptor. In particular, it refers to a polypeptide which has an immunomodulatory activity. In one aspect of the invention, a polypeptide of the invention may activate NFkB or may elicit pro-inflammatory cytokine production or induce expression of cell surface co-stimulatory receptors required for activation of T-cells.

15

Alternatively, a variant of the polypeptide of the invention is one which exhibits binding to the same ligand as TLR9. Such ligand binding may be assayed using the assays described below.

20

In other aspects of the invention a variant is one which does not show the same function as TLR9 but which may be used to inhibit this function. For example, a variant polypeptide for use in an assay or therapy is one which inhibits TLR9 activity, for example by inhibiting or compelling out ligand binding or receptor complex formation by TLR9. Alternatively, a variant may be one which inhibits ligand binding to TLR9. Such a variant may inhibit activation of NFkB or inhibit cytokine production and expression of cell surface co-stimulatory receptors.

25

Such inhibitors may be used as immunomodulators to reduce inappropriate TLR activation in asthma or other chronic inflammatory diseases, or septic shock.

5

To determine whether a variant has the same essential function as TLR9, the immunomodulatory activity can be determined by monitoring the effect of a substance on different immune responses. For example the effect of the substance under test on NFkB activation mediated through binding the polypeptide of the present invention may be monitored. This can be carried out, for example, by co-transfection of a construct expressing the polypeptide with a construct containing a reporter gene, such as secreted placental alkaline phosphatase, under the control of a suitable NFkB-responsive promoter and monitoring for expression of the reporter gene.

10

Alternatively, other immunomodulatory activity such as the production of cytokines can be determined by monitoring cytokine production following incubation of a test substance with a cell expressing a polypeptide of the invention. Such assays may be carried out in the presence or absence of additional T-lymphocytes to assess the effect of such cytokines, or the direct action of a polypeptide of the invention, on such T-lymphocytes to thus determine immunomodulatory activity.

15

20

Alternatively, the Toll-like receptor functionality is as a peptide which binds a ligand of TLR9, inhibits immunomodulatory activity by TLR9 or inhibits ligand binding to TLR9 and can be determined by an assay as described below.

25

Preferably, a polypeptide of the invention will show the structural features associated with a Toll-like receptor. Preferably, a polypeptide of the invention, or a functional fragment thereof, contains one or more of the following structural features associated with a Toll-like receptor: an extracellular region containing leucine-rich repeat motif and cysteine-rich regions involved in ligand binding; a

30

single hydrophobic transmembrane region; and a C-terminal cytoplasmic domain sharing homology with other TLRs and members of the IL-1 receptor family.

Typically, polypeptides with more than about 65% identity, preferably at least 80% or at least 90% and particularly preferably at least 95%, at least 97%, or at least 99% identity, with the amino acid sequences of SEQ ID NO: 2 over a region of at least 20, preferably at least 30, at least 40, at least 60 or at least 100 contiguous amino acids or over the full length of SEQ ID NO: 2, are considered as variants of the proteins. Identity is calculated using the widely used GCG (University of Wisconsin) suite of programs and preferably using the distances software (correction method). Such variants may include allelic variants and the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains the basic biological functionality of the Toll-like receptor, having a similar function to TLR9 or inhibits such function such as preventing ligand binding or TLR9 mediated activation. Such variants also include isoforms such as TLR9-A, which is 23 amino acids (or 2.2%) shorter than TLR9 (see SEQ ID NO: 2 and SEQ ID NO: 4) and thus shows 97.8 identity therewith. Transcription of the nucleotide sequence presented in SEQ ID NO: 1 can result, due to variable mRNA splicing involving a second exon encoding an alternative initiating methionine, in an mRNA having the sequence of SEQ ID NO: 3 which, when translated, results in the polypeptide TLR9-A depicted in SEQ ID NO: 4.

Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions. The modified polypeptide generally retains activity as a TLR9 receptor or inhibitor of TLR9 receptor activity. Conservative substitutions may be made, for example according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

ALIPHATIC	Non-polar	G A P	
		I L V	
	Polar-uncharged	C S T M	
	Polar-charged	N Q	
AROMATIC		D E	
		K R	
		H F W Y	

Shorter polypeptide sequences are within the scope of the invention. For example, a peptide of at least 20 amino acids or up to 50, 60, 70, 80, 100 or 150 amino acids in length is considered to fall within the scope of the invention as long as it demonstrates the basic biological functionality of TLR9 or inhibits TLR9. In accordance with this aspect of the invention the peptide may also comprise an epitope of TLR9 for generation of antibodies. In particular, but not exclusively, this aspect of the invention encompasses the situation when the protein is a fragment of the complete protein sequence and may represent a ligand-binding region (N-terminal extracellular domain) or an effector binding region (C-terminal intracellular domain). Fragments from which the C-terminus has been removed may be used as decoy receptors. Other fragments such as a secreted or soluble form of the receptor may be generated for use in an assay or in therapy in accordance with the invention. Such fragments can also be used to raise anti-TLR9 antibodies.

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or may comprise modified amino acid residues. They may also be modified by the addition of histidine residues or an epitope tag for example by a (His) 8 or (His) 6 sequence or a HA, TT, Myc or Flag tag to assist their purification or detection. They may be modified by the addition of a signal sequence to promote insertion into the

cell membrane. Such modified polypeptides fall within the scope of the term "polypeptide" of the invention.

The invention also includes nucleotide sequences that encode for TLR9 or variants thereof as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA including genomic DNA, synthetic DNA or cDNA. Preferably the nucleotide sequence is a DNA sequence and most preferably, a cDNA sequence. Nucleotide sequence information is provided in SEQ ID NO: 1. Such nucleotides can be isolated from human cells or synthesised according to methods well known in the art, as described by way of example in Sambrook *et al.* Such nucleotides can typically be isolated from activated cells of the immune system, heart, lung, pancreatic islet cells and lymph nodes, adenoid and tonsil tissues. Figures 2 and 3 show the tissue distribution of RNA encoding TLR9, as determined by extraction of total RNA from normal or disease tissue or cells which is then used to generate cDNA for real time quantitative PCR using suitable primers and probes (TaqMan analysis) to assess expression patterns. The profiles show differential expression across tissues tested and predominance to sites containing inflammatory cell types.

Typically a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridising under selective conditions to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1.

A polynucleotide of the invention can hybridize to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1 (or of SEQ ID NO: 3) at a level significantly above background. Background hybridisation may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the coding sequence or complement of the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3 is typically at least 10 fold, preferably at least 100 fold, as intense

as interactions between other polynucleotides and the coding sequence of SEQ ID NO: 1 or of SEQ ID NO: 3. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ^{32}P . Selective hybridisation may typically be achieved using conditions of low stringency (0.03M sodium chloride and 0.03M sodium citrate at about 40°C), medium stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 50°C) or high stringency (for example, 0.03M sodium chloride and 0.03M sodium citrate at about 60°C).

The coding sequence of SEQ ID NO: 1 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 25, 50 or 100 substitutions. The polynucleotides of the present invention may alternatively or additionally be modified by one or more insertions and/or deletions and/or by an extension at either or both ends. The modified polynucleotide generally encodes a polypeptide which has Toll-like receptor activity or inhibits the activity of TLR9. Degenerate substitutions may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in the Table above.

A nucleotide sequence of the invention which is capable of selectively hybridising to the complement of the DNA coding sequence of SEQ ID NO: 1 will generally have at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, more preferably at least 100 contiguous nucleotides or most preferably over the full length of SEQ ID NO: 1. Methods of measuring nucleic acid and protein homology are well known in the art. For example the UWGCG Package provides the BESTFIT program which can be used to calculate homology (Devereux *et al* 1984). Similarly the PILEUP and BLAST algorithms can be used to line up sequences (for example are described in Altschul 1993,

and Altschul *et al* (1990). Many different settings are possible for such programs. In accordance with the invention, the default settings may be used.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher sequence identity over longer lengths) being preferred. Thus, for example a polynucleotide which has at least 80% sequence identity over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which has at least 95% sequence identity over 40 nucleotides. The most preferred sequences have at least 70% sequence identity over at least 70% of the full length of the sequence provided by SEQ ID NO: 1.

The nucleotides according to the invention have utility in production of the proteins according to the invention, which may take place *in vitro*, *in vivo* or *ex vivo*. The nucleotides may be involved in recombinant protein synthesis or indeed as therapeutic agents in their own right, utilised in gene therapy techniques. Nucleotides complementary to those encoding TLR9, or antisense sequences, may also be used in gene therapy, such as in strategies for down regulation of expression of the proteins of the invention.

Polynucleotides of the invention may be used as a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will preferably be at least 10, preferably at least 15 or at least 20, for example at least 25, at least 30 or at least 40 nucleotides in length. They will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150

nucleotides in length, for example up to 200, 300, 400, 500 nucleotides in length, or even up to a few nucleotides, such as five or ten nucleotides, short of the coding sequence of SEQ ID NO: 1.

The present invention also includes expression vectors that comprise nucleotide sequences encoding the proteins or variants thereof of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to a person skilled in the art. By way of further example in this regard we refer to Sambrook *et al*.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used as test compounds in the assays of the invention or may be useful in a method of treatment of the human or animal body by therapy.

Preferably, a polynucleotide of the invention or for use in the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence, such as a promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example in a method of gene therapy.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which expression is designed. For example, yeast promoters include *S. cerevisiae* GAL4 and ADH promoters, *S. pombe* *nmt1* and *adh* promoter. Mammalian promoters include the metallothionein promoter which can be induced in response to heavy metals such as cadmium. Viral promoters such as the SV40 large T antigen promoter or adenovirus promoters may also be used. All these promoters are readily available in the art.

Mammalian promoters, such as β -actin promoters, may be used. Tissue-specific promoters may be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). Viral promoters are readily available in the art.

The vector may further include sequences flanking the polynucleotide which comprise sequences homologous to eukaryotic genomic sequences, preferably mammalian genomic sequences, or viral genomic sequences. This will allow the

introduction of the polynucleotides of the invention into the genome of eukaryotic cells or viruses by homologous recombination. In particular, a plasmid vector comprising the expression cassette flanked by viral sequences can be used to prepare a viral vector suitable for delivering the polynucleotides of the invention to a mammalian cell. Other examples of suitable viral vectors include herpes simplex viral vectors and retroviruses, including lentiviruses, adenoviruses, adeno-associated viruses and HPV viruses (such as HPV-16 or HPV-18). Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for example may be used to stably integrate the polynucleotide giving rise to the RNA into the host genome. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression.

The invention also includes cells that have been modified to express the TLR9 polypeptide or a variant thereof. Such cells include transient, or preferably stable higher eukaryotic cell lines, such as mammalian cells or insect cells, lower eukaryotic cells, such as yeast or prokaryotic cells such as bacterial cells. Particular examples of cells which may be modified by insertion of vectors encoding for a polypeptide according to the invention include mammalian HEK293T, CHO, HeLa and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of a polypeptide. Cells such as T-cells, monocytes or dendritic cells expressing the receptor may be used for example in screening. Expression may be achieved in transformed oocytes. A polypeptide of the invention may be expressed in cells of a transgenic non-human animal, preferably a mouse. A transgenic non-human animal expressing a polypeptide of the invention is included within the scope of the invention.

It is also possible for the proteins of the invention to be transiently expressed in a cell line or on a membrane, such as for example in a baculovirus expression

18

system. Such systems, which are adapted to express the proteins according to the invention, are also included within the scope of the present invention.

5 According to another aspect, the present invention also relates to antibodies (either polyclonal or preferably monoclonal antibodies, chimeric, single chain, Fab fragments) which have been raised by standard techniques and are specific for a polypeptide of the invention. Such antibodies could for example, be useful in purification, isolation or screening methods involving immunoprecipitation techniques and may be used as tools to further elucidate the function of TLR9 or a variant thereof, or indeed as therapeutic agents in their own right. Antibodies may also be raised against specific epitopes of the proteins according to the invention. Such antibodies may be used to block ligand binding to the receptor. Alternatively an antibody may be provided which acts as an agonist, to cross link receptors of the invention to mediate receptor activity. An antibody, or other compound, "specifically binds" to a protein when it binds with high affinity to the protein for which it is specific but does not bind or binds with only low affinity to other proteins. A variety of protocols for competitive binding or immunoradiometric assays to determine the specific binding capability of an antibody are well known in the art (see for example Maddox *et al* 1993). Such immunosays typically involve the formation of complexes between the "specific protein" and its antibody and the measurement of complex formation.

20 An important aspect of the present invention is the use of polypeptides according to the invention in screening methods to identify compounds that may act as modulators of Toll-like receptor activity. Any suitable form may be used for the assay to identify a modulator of TLR9 activity. In general terms, such screening methods may involve contacting a polypeptide of the invention with a test compound and then measuring receptor activity.

30

19

Screening methods may alternatively involve contacting a polypeptide of the invention with a test compound and then monitoring for the effect on immunomodulatory activity.

5 The binding of the substance to a polypeptide in the invention can be determined directly. For example, a radiolabelled test substance can be incubated with a polypeptide of the invention and so that binding of the test substance to the polypeptide can be monitored. Typically, the radiolabelled test substance can be incubated with cell membranes containing the polypeptide until equilibrium is reached. The membranes can then be separated from a non-bound test substance and dissolved in scintillation fluid to allow the radioactive content to be determined by scintillation counting. Non-specific binding of the test substance may also be determined by repeating the experiments in the presence of a saturating concentration of a non-radioactive ligand. Preferably, a binding curve is constructed by repeating the experiment with various concentrations of the test substance.

15 Cell based assays may also be carried out, for example using a cell expressing the TLR9 receptor, and contacting the cell with another cell to look for ligand binding or activation of TLR9-mediated pathways such as NF κ B activation. Alternatively cells expressing TLR9 constitutively may be provided for use in assays for TLR9 function. Such constitutively expressed TLR9 may demonstrate TLR9 activity in the absence of ligand binding. Additional test substances may be introduced in any assay to look for inhibitors of ligand binding or inhibitors of TLR9-mediated activity. Assays are preferably carried out using cells expressing TLR9, and incubating such cells with the test substance optionally in the presence of TLR9 ligand. Alternatively an antibody may be used to complex TLR9 and thus mediate TLR9-activity. Test substances may then be added to assess the effect on such activity.

In preferred aspects, a host cell is provided expressing the receptor, or a receptor complex of TLR8 (or TLR8-A) comprising a homodimer, a heterodimer of TLR8 (or TLR8-A) with another Toll-like receptor, or a complex of TLR9 (or TLR9-A) with protein cofactors, and containing an NFkB responsive reporter construct. The host cell is treated with a substance under test for a defined time. The expression of the reporter gene, such as secreted placental (SP) alkaline phosphatase or luciferase is assayed. The assay enables determination of whether the addition of compounds inhibits the induction of the response in target cells. Alternatively the assay may be carried out to identify cytokine production or it may be carried out in the presence of T-cells to identify inducement of co-stimulatory receptors required for activation of T-cells.

Assays may also be carried out to identify modulators of receptor shedding. A polypeptide of the invention can be cleaved from the cell surface. Shedding the receptor would act to down regulate receptor signalling. Thus, cell based assays may be used to screen for compounds which promote or inhibit receptor-shedding. Assays may also be carried out to identify substances which modify TLR8 receptor expression for example substances which down regulate expression. Such assays may be carried out for example by using antibodies for TLR8 to monitor levels of TLR8 expression.

Additional control experiments may be carried out. Assays may also be carried out using known ligands of other Toll-like receptors to identify ligands which are specific for polypeptides of the invention. Preferably, the assays of the invention are carried out under conditions which would result in Immunomodulatory NFkB mediated activity in the absence of the test substance, to identify inhibitors of Toll-like receptor mediated activity, or agents which inhibit ligand-induced Toll-like receptor activity.

Suitable test substances which can be tested in the above assays include combinatorial libraries, defined chemical entities, peptides and peptide mimetics, oligonucleotides and natural product libraries, such as display (e.g. phage display libraries) and antibody products. In a preferred embodiment, the test substance is a variant peptide of the invention. In particularly preferred embodiments, suitable test substances which may be candidate ligands for binding to and modulation of TLR8 or TLR8-A include synthetic or naturally occurring oligonucleotides or polynucleotides which bind directly to the Toll-like receptor or which modulate the Toll-like receptor of the present invention via one or more intermediate signal transducers. Such oligo- or polynucleotides may be "CG-rich" sequences or sequences including one or more unmethylated CpG nucleotide pairs.

The assay may be carried out using full length TLR8 to identify a variant peptide which interferes with TLR8 mediated activity, for example by inhibiting ligand binding.

Test substances may be used in an initial screen of, for example, 10 substances per reaction, and the substances of these batches which show inhibition or activation tested individually. Test substances may be used at a concentration of from 1nM to 1000µM, preferably from 1µM to 100µM, more preferably from 1µM to 10µM.

A protein-binding assay may be developed using a polypeptide of the invention, preferably one encoding the extracellular ligand-binding domain, to identify novel protein ligands of TLR8. Particular examples may be screening of a human cDNA expression library for protein ligands of TLR8 by yeast 2-hybrid protein interaction assay (e.g. as described in International Patent Application No. WO98/49294).

22

Another aspect of the present invention is the use of polynucleotides encoding the TLR9 polypeptides of the invention to identify mutations in TLR9 genes which may be implicated in human disorders or to identify cells in which TLR9 is expressed. Identification of such mutations may be used to assist in diagnosis of immune system, lung, kidney, heart or other disorders or susceptibility to such disorders and in assessing the physiology of such disorders. In particular the polynucleotides of the invention may assist in diagnosis of asthma and rheumatoid arthritis. For example, a SNP (single nucleotide polymorphism) has been identified in the genomic DNA encoding TLR9 (G/A nucleotide: The SNP Consortium database accession number TSC0164834). This single base pair change lies in the DNA region encoding the 23 N-terminal residues of TLR9, and this region is spliced out of the mRNA encoding TLR9-A. The nucleotide at this SNP position may affect the efficiency of mRNA splicing in the two different variants - a G at this position may possibly disrupt the splicing machinery and an A might lead to more efficient splicing. Additionally, the presence of a G as compared to an A in an unspliced mRNA would introduce a stop codon and result in different N terminal protein sequences upon translation of that mRNA, thus the two polymorphic variants of the *tlr9* gene may encode receptors which have differing expression levels and/or differing functional activity levels. The present invention provides a diagnostic tool for determining the polymorphic variant in an individual by detecting the DNA sequence at the SNP site. Such a tool may incorporate a nucleotide probe specific for one or other of the polymorphic variants, for example an oligonucleotide of from 5 to 50, preferably 5-20 nucleotides, complementary to a fragment of the nucleotide sequence of SEQ ID No. 1 which extends over the SNP site or a fragment complementary to that sequence with the exception of the single nucleotide change (G to A) at the SNP site. The present invention also provides a method of detecting a polymorphic variant in the *tlr9* gene by determining the sequence of nucleotides at and around the SNP site identified by the SNP consortium database accession number

23

TSC0164834, in particular by determining whether the nucleotide at that SNP site is a G or an A.

Another aspect of the present invention is the use of the compounds that have been identified by screening techniques referred to above in the treatment or prophylaxis of disorders which are responsive to regulation of TLR9 receptor activity. In addition, variant peptides of the invention which inhibit TLR9-mediated activity, for example which inhibit ligand binding or prevent hTLR9 immunomodulatory activity may be used in the treatment or prophylaxis of such disorders. Antibodies which recognise TLR9 may similarly be used in therapy.

In particular, such compounds may be used in the treatment of inflammatory, cardiovascular, systemic infection or autoimmune disease. The compounds may be used to treat bacterial, viral or fungal infections, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis and Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, septic shock syndrome associated with systemic infection involving gram positive and gram negative bacteria, diabetes, Multiple Sclerosis.

In an alternative aspect, the invention provides agents which activate TLR9 mediated immunomodulation for use as an immunoadjuvant, or TLR9, and variants thereof, or polynucleotides or DNA encoding a polypeptide of the invention which may be administered for use as immunoadjuvants to enhance or alter the immune response in an individual to an antigen.

The compounds identified according to the screening methods outlined above may be formulated with standard pharmaceutically acceptable carriers and/or excipients as is routine in the pharmaceutical art, and as fully described in Remington's Pharmaceutical Sciences, Mack Publishing Company, Eastern

Pennsylvania 17th Ed. 1985, the disclosure of which is included herein of its entirety by way of reference.

5 The compounds may be administered by enteral or parenteral routes such as via oral, buccal, anal, pulmonary, intravenous, intra-arterial, intramuscular, intraperitoneal, topical or other appropriate administration routes. The dose of a compound to be administered may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the patient to be treated; the route of administration; and the required regimen. Again, a physician will be able to determine the required route of administration and dosage for any particular patient. A typical daily dose is from about 0.1 to 50 mg per kg of body weight, according to the activity of the compound, the age, weight and conditions of the subject to be treated, the type and severity of the disease and the frequency and route of administration. Preferably, daily dosage levels are from 5 mg to 2 g.

15 Nucleic acid encoding an inhibitor of TLR9 activity may be administered to the mammal. In an alternative aspect of the invention, nucleic acid encoding TLR9 or a variant thereof may be administered to the animal. Such a variant shows immunomodulatory activity of TLR9 such as inducing cytokine production and expression of cell surface co-stimulatory receptors required for activation of T-cells. Nucleic acid, such as RNA or DNA, and preferably, DNA, is provided in the form of a vector, such as the polynucleotides described above, which may be expressed in the cells of the mammal.

25 Nucleic acid encoding the peptide may be administered to the animal by any available technique. For example, the nucleic acid may be introduced by injection, preferably intradermally, subcutaneously or intramuscularly. Alternatively, the nucleic acid may be delivered directly across the skin using a nucleic acid delivery device such as particle-mediated gene delivery. The

nucleic acid may be administered topically to the skin, or to the mucosal surfaces for example by intranasal, oral, intravaginal, intrarectal administration.

5 Uptake of nucleic acid constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents includes cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the nucleic acid to be administered can be altered. Typically the nucleic acid is administered in the range of 1pg to 1mg, preferably to 1pg to 10µg nucleic acid for particle mediated gene delivery and 10µg to 1mg for other routes.

Examples

15 Example 1 - Screening for compounds which exhibit protein modulating activity

Mammalian cells, such as HeK293, CHO and HeLa cells over-expressing TLR9 protein are generated for use in the assay. 96 and 384 well plate, high throughput screens (HTS) are employed. TLR9 induced cell activation can be monitored through activation of a signal transduction event (typically resulting in activation of NFκB or AP-1) or transcriptional activation of a reporter gene (typically regulated via NFκB or AP-1 responsive elements). TLR9 induced activation of a reporter gene such alkaline phosphatase or luciferase can be easily assessed using a suitable colourimetric or fluorimetric assay to measuring production. Such assays may be used to identify receptor antagonists capable of blocking ligand induced TLR9 activation, inhibitors capable of blocking TLR9 intracellular signal transduction or receptor agonists capable of activating TLR9. Secondary screening involves evaluation of TLR9 modulators in disease related

28

cells. Tertiary screens involve the study of modulators in rat and mouse models of disease relevant to the target.

27

SEQUENCE LISTING

5 <110> GLAXO GROUP LIMITED
<120> PROTEIN
<130> P78721 GCM SER
10 <140>
<141>
<150> 2
<170> Patentin Ver. 2.1
15 <210> 1
<211> 3165
<212> DNA
<213> Homo sapiens
20 <220>
<221> CDS
<222> (1)..(3165)
25 <400> 1
atg ccc atg aag tgg agt ggg tgg agg tgg agc tgg ggg ccg gcc act 48
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr
1 5 10 15
cac aca gcc ctc cca ccc cca cag ggt ttc tgc cgc agc gcc ctg cac 96
30 His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His
20 25 30
ccg ctg tct ctc ctg gtc cag gcc atc atg ctg gcc atg acc ctg gcc 144
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
ctg ggt acc ttg cct gcc ttc cta ccc tgt aag ctg cag ccc cac ggc 192
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Gln Pro His Gly
50 55 60
ctg gtc aac tgc aac tgg ctg ttc ctg aag tct gtc ccc cac ttc tcc 240
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
65 70 75 80
atg gca gca ccc ggt gcc aat gtc acc agc ctt tcc ttg tcc tcc aac 288
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
85 90 95
cgc atc cac cac ctc cat ggt tct gac ttt gcc cac ctg ccc agc ctg 336
45 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
cgg cat ctc aac ctc aag tgg aac tgc ccg ccg gtt gcc ctg agc ccc 384
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
atg cac ttc ccc tgc cac atg acc atc gag ccc agc acc ttc ttg ggt 432
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
130 135 140

480 gtr ccc acc ctg gaa gaa cta aac aac ctg agc tac aac aac atc atg act
 Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
 145 150 155 160
 528 gtr cct ggg ctg ccc aaa tcc ctc cta ctg tcc ctc agc cat acc
 5 Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
 165 170 175 178
 576 aac atc ctg atg cta gac tct ggc ctc ggc ctg cat ggc ctg
 Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
 180 185 190 195
 624 10 cgc ttc cta ttc atg gac ggc aac tct tat tac aag aac ccc tgc agt
 Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
 195 200 205 210
 672 cag gaa ctg ggg ggc ccc ggt ggc ctc ctc ggc ctg ggc aac ctc
 Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
 210 215 220 225
 720 aac ccc ctg tca ctc aag tac aac aac ctc act gtr gtr ccc ggc aac
 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Pro Arg Asn
 225 230 235 240
 768 ctg cct tcc agc ctg gaa tat ctg ctg tta tta tta aac cgc atc gtc
 Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val
 245 250 255 260
 816 aaa ctg gtr cct ggc gac ctg ggc aac ttc acc ggc ctg gtr gtr ctc
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Thr Thr Ala Leu Arg Val Leu
 260 265 270 275
 864 gtr gtr ggc ggc aac tgc ggc gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asp Val Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
 275 280 285 290
 912 atg gtr gtr cct gtr ctc ttc ccc ccc ccc ccc ccc ccc ccc ccc ccc
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser
 295 300 305 310
 960 cgc ctg gtr cct ctt gaa gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser
 310 315 320 325
 1008 tgc tgc aat gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
 325 330 335 340
 1056 gtr ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
 340 345 350 355
 1104 ttc cag gtr cta aca cag ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
 355 360 365 370
 1152 cca aag aag gtr ttc ttc ggc ctc ctc ctc ctc ctc ctc ctc ctc ctc
 Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Leu Ala Pro Ser Phe Gly
 370 375 380 385
 1200 gtr ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
 385 390 395 400
 1248 tgc ctc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 50 Ser Leu Asp Glu Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
 405 410 415 420
 1296 cgc act ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr

Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
 420 425 430 435
 1344 atc ttc agt gtr ctc ctc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
 440 445 450 455
 1392 cgc atc agt gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Arg Ile Ser Gly Ala Ser Glu Leu Thr Thr Thr Thr Thr Thr Thr Thr
 460 465 470 475
 1440 gaa gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Gly Gly Gly Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
 480 485 490 495
 1488 gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Val Asp Thr Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 500 505 510 515
 1536 aac ttc acc ttr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
 520 525 530 535
 1584 gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Gln Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
 540 545 550 555
 1632 aac tgc atc tgc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
 560 565 570 575
 1680 gtr ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
 580 585 590 595
 1728 gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Gln His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
 600 605 610 615
 1776 ttc aac agt gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
 620 625 630 635
 1824 ttc gtr gtr ctc ctc gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
 640 645 645 650
 1872 aac atc ctc agt gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
 650 655 660 665
 1920 gtr ctg gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Gly
 665 670 675 680
 1968 gtr ctc tat ctg ctc ttc ttc ttc ttc ttc ttc ttc ttc ttc ttc ttc
 Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
 680 685 690 695
 2016 gtr ttr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
 700 705 710 715
 2064 aac ctc ccc aag agt cta cag gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
 720 725 730 735
 2112 gtr ttc ttc aag ttr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr gtr
 Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val

30

690 ctc gac ctg gca gga aac cag ctg aag gcc ctg acc aac gcc ctg 2160
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
705 710 715 720
5 cct gct gcc acc cpg ctg cag aag ctg gat gtc agc tpc aac gcc atc 2208
Pro Ala Gly Thr Arg Leu Arg Arg Leu Val Ser Cys Asn Ser Ile
725 730 735
agc ttc gta gcc ccc gcc ttc ttc acc aag gcc aag gac ctg cag gaa 2256
Ser Phe Val Ala Pro Gly Phe Ser Lys Ala Lys Gln Leu Arg Gln
740 745 750
ctc aac ctt agc gcc aac gcc ccc aag acta gta gac cac ttc tgg ttc 2304
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
755 760 765
ggg ccc ctg gga agt gcc ctg caa ata cta gat gta agc gcc aac cct 2352
15 Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
770 775 780
ctc cac tgc gcc tct gaa gaa gcc ttt atg gcc ttc ctg ctg gaa gta 2400
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Gln Val
785 790 795
20 cag gct gcc gta gcc gct cag ccc agc cpg gta aag tct gcc agt cag 2448
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
805 810 815
ggc cag ctg cag gcc ctg gcc atc ttt gca cag gac ctg gcc ctg tgc 2496
Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu Cys
820 825 830
ctg gat gaa gcc ctg ttc tgg gac tct ttc gcc ctg tgc ctg ctg gct 2544
Leu Asp Gln Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
835 840 845
gta gct ctg gcc ctg gta gta cct atg ctg cat cac ctg tct gcc tgg 2592
30 Val Ala Leu Gly Leu Gly Val Pro Met Leu His Lys Cys Gly Trp
850 855 860
ggc ctg tgg tac tgc ttc cac ctg tgc ctg gcc tgg ctt ccc tgg cag 2640
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
865 870 875
35 ggg cpg caa agt ggg cga gat gaa gac gcc ctg ccc tac gat gcc ttc 2688
Gly Arg Gln Ser Gly Arg Asp Gln Asp Ala Leu Pro Tyr Asp Ala Phe
885 890 895
gta gcc ttc gac aac aag cag agc gca gta gac tgg gaa tac aac 2736
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn
900 905 910
40 gaa ctt cpg gaa cag ctg gaa gaa tgc cgt gaa gcc tgg paa ctc cgc 2784
Gln Leu Arg Gly Gln Leu Gln Cys Arg Gly Arg Trp Ala Leu Arg
915 920 925
ctg tgc ctg gaa gaa gcc gac tgg ctg cct gcc aac acc ctc ttt gaa 2832
45 Leu Cys Leu Gln Gly Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Gln
930 935 940
aac ctg tgg gcc tgg gtc tct gcc agc cgc aag aag ctg ttt gta ctg 2880
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
945 950 955
50 gac cac aag gac cag gcc agt gct ctg ctg gcc gcc agc ttc ctg ctg 2928
Ala His Thr Asp Arg Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
965 970 975

31

ggc cag cag gcc ctg ctg gaa gcc cgc aag gcc gta gta ctg ctg gta 2976
Ala Gln Gln Arg Leu Leu Gln Asp Arg Lys Asp Val Val Leu Val
980 985 990
atc ctg agc cct gac gcc cgc cgc ttc cgc tat gta cpg ctg gcc cag 3024
5 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln
995 1000 1005
cgc ctg tgc cgc cag agt gta ctg ctc ttc tgg ccc cac cag ccc agt ggt 3072
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
1010 1015 1020
10 cag gcc agc ttc tgg gcc cag ctg gcc atg gcc ctg acc aag gcc aac 3120
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
1025 1030 1035
cac cac ttc tat aac cpg aac ttc tgc cag gga ccc acg gcc gaa 3165
15 His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Gln
1045 1050 1055
<210> 2
<211> 1055
<212> PRT
<213> Homo sapiens
20
<400> 2
Met Pro Met Lys Trp Ser Gly Trp Arg Trp Ser Trp Gly Pro Ala Thr
1 5 10 15
25 His Thr Ala Leu Pro Pro Gln Gly Phe Cys Arg Ser Ala Leu His
20 25 30
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
35 40 45
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Gln Leu Gln Pro His Gly
50 55 60
30 Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
65 70 75
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
80 85 90 95
35 Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
100 105 110
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
115 120 125
Met His Phe Pro Cys His Met Thr Ile Gln Pro Ser Thr Phe Leu Ala
130 135 140
40 Val Pro Thr Leu Gln Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
145 150 155
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
160 165 170 175
45 Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
180 185 190
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
195 200 205
Gln Ala Leu Gln Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
210 215 220 225
50 Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
225 230 235 240

32

Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val 245
 Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu 250
 5 Asp Val Gly Gly Asn Cys Arg Cys Asp His Ala Pro Asn Pro Cys 255
 Met Glu Cys Pro Arg His Phe Pro Glu Leu His Pro Asp Thr Phe Ser 260
 His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser 265
 10 Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu 270
 Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala 275
 15 Phe Glu Gly Leu Thr Glu Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr 280
 Glu Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly 285
 Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg 290
 20 Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu 295
 Glu Thr Leu Arg Leu Glu Met Asn Phe Ile Asn Glu Ala Glu Leu Gly 300
 25 Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn 305
 Arg Ile Ser Gly Ala Ser Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp 310
 30 Gly Gly Glu Lys Val Trp Leu Glu Pro Gly Asp Leu Ala Pro Ala Pro 315
 Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu 320
 Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Glu Pro 325
 35 Glu Met Phe Ala Glu Leu Ser His Leu Glu Cys Leu Arg Leu Ser His 330
 Asn Cys Ile Ser Glu Ala Val Asn Gly Ser Glu Phe Leu Pro Leu Thr 335
 Gly Leu Glu Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His 340
 40 Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser 345
 Tyr Asn Ser Glu Pro Phe Gly Met Glu Gly Val Gly His Asn Phe Ser 350
 45 Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn 355
 Asn Ile His Ser Glu Val Ser Glu Glu Leu Cys Ser Thr Ser Leu Arg 360
 50 Ala Leu Asp Phe Ser Gly Asn Ala Leu Glu Gly His Met Trp Ala Glu Gly 365
 Asp Leu Tyr Leu His Phe Phe Glu Gly Leu Ser Gly Leu Ile Trp Leu 370

EMBOSS: vnc_01808061.1

33

Asp Leu Ser Glu Asn Arg Leu His Thr Leu Leu Pro Glu Thr Leu Arg 665
 Asn Leu Pro Lys Ser Leu Glu Val Leu Arg Leu Arg Asp Asn Tyr Leu 670
 5 Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val 675
 Leu Asp Leu Ala Gly Asn Glu Leu Lys Ala Leu Thr Asn Gly Ser Leu 680
 705 Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile 700
 Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu 705
 10 Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe 710
 15 Gly Pro Leu Ala Ser Ala Leu Glu Ile Leu Asp Val Ser Ala Asn Pro 715
 Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Glu Val 720
 785 Glu Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro 725
 20 Gly Glu Leu Glu Gly Leu Ser Ile Phe Ala Glu Asp Leu Arg Leu Cys 730
 Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Ala 735
 25 Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp 740
 Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg 745
 865 Gly Arg Glu Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe 750
 30 Val Val Phe Asp Lys Thr Glu Ser Ala Val Ala Asp Trp Val Tyr Asn 755
 Glu Leu Arg Gly Glu Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 760
 35 Leu Cys Leu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu 765
 Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu 770
 945 Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu 775
 40 Ala Glu Glu Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val 780
 Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Glu 785
 45 Arg Leu Cys Arg Glu Ser Val Leu Leu Trp Pro His Glu Pro Ser Gly 790
 Glu Arg Ser Phe Trp Ala Glu Leu Gly Met Ala Leu Thr Arg Asp Asn 795
 1025 His His Phe Tyr Asn Arg Phe Cys Glu Gly Pro Thr Ala Glu 800
 50 1035 1040 1050 1055

EMBOSS: vnc_01808061.1

34

Seq.id.no.3

atg ggt ttc tgc cgc agc gcc ctg cac
met gty phe cys arg ser ala leu his
5
ccg ctg tct ctc ctg ctg cag gcc atc atg ctg gcc atg acc ctg gcc 75
Pro Leu Ser Leu Leu Val Gln Ala Ile Met Leu Ala Met Thr Leu Ala
10
ctg ggt acc ttg cct gcc ttc cta ccc tgt gag ctc cag ccc cag gcc 123
Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Gly Leu Gln Pro His Gly
15
ctg ggt aac tgc aac tgg ctg ttc ctg aag tct gtg ccc cag ttc ttc 171
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser
20
atg gca gca cct cgt gcc aat gtc acc agc ctt tcc ttg tcc tcc aac 219
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
25
cgc atc cag cag ctc cat gat tct gac ttt gcc cag ctg ccc gcc ctg 267
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
30
cgg cat ctc aac ctg aag tgg aac tgc cgg cgg gtt gcc ctg acc ccc 315
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
35
atg cac ttc ccc tgc cag atc atg acc atc gag ccc agc acc ttc ttg gct 363
Met His Phe Pro Cys His Met Thr Ile Gly Pro Ser Thr Phe Leu Ala
40
gtg ccc acc ctg gaa gag cta aac ctg agc tac aac aac atc atg act 411
Val Pro Thr Leu Gly Leu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
45
gtg cct ggg ctg ccc aaa tcc ctc tta tcc ctg tcc ctc agc cat acc 459
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
50
aac atc ctg atg cta gcc tct gcc agc ctc gcc ctg ctg cat gcc ctg 507
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
55
cgc ttc cta ttc atg gac gcc aac tgt tat tac aag aac ccc tgc ggg 555
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
60
cag gca ctg gag gtc gcc cgg gtc ctc ctt gcc ctg ggc aac ctc 603
Gln Ala Leu Gly Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
65
acc cag ctg tca ctc aag tac aac acc act atg gtg gtg ccc ggc aac 651
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
70
ctg cct ttc agc ctg ggg tat ctg ctg ttg tcc tac aac gcc atc gtc 699
Leu Pro Ser Ser Leu Gly Tyr Leu Leu Ser Tyr Asn Arg Ile Val
75
aaa ctg ggg cct gag gcc ctg gcc aat ctg acc gcc ctg ggt gtc 747
Lys Leu Ala Pro Gly Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
80
ggt gtg gcc gga aat tgc gcc cgt tgc gcc cag cct ccc aac ccc tgc 795
Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
250 255 260 265

Seq.id.no.4

Seq.id.no.5

35

atg gag tgc cct cgt cac ttc ccc cag cta cat ccc gtt acc ttc agc 843
Met Gly Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
890
cac ctg agc cgt cct gaa gcc ctg ttg ttg aag agt tct ctc tcc 891
His Leu Ser Arg Leu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser
940
tgg ctg aat gcc agt tgg ttc cgt ggg ctg gga aac ctc cga ggt ctg 939
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Arg Val Leu
990
gac ctg agt gag aac ttc ctc tac aaa tgc atc act aaa acc aag gcc 987
Asp Leu Ser Gly Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
1040
ttc cag gcc cta acc cag ctg cgc aag ctt aac ctg ttc ttc aat tac 1035
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
1090
caa aag agg ttg ttc ttc gcc cag ctg tct ctg gcc cct tcc ttc ggg 1083
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
1140
agc ctg gtc gcc ctg aag gag ctg gcc atc cag gcc atc ttc ttc cgc 1131
Ser Leu Val Ala Leu Lys Gly Leu Asp Met His Gly Ile Phe Phe Arg
1190
tea ctc gat gag acc aag ctc cgg cca ctg gcc cgc ctg ccc atg ctc 1179
Ser Leu Asp Gly Thr Thr Leu Arg Pro Leu Ala Arg Leu Met Leu
1240
cag act ctg cgt ctg cag atg aac ttc atc aac cag gcc cag ctg gcc 1127
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
1290
atc ttc agg gcc ttc cct gcc ctg cgc tac ttg gtc ctg tgc gac aac 1275
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
1320
ctg acc gcc gga gtt tgc gag ctg aca gcc acc atc ggg gag gca gtt 1323
Arg Ile Ser Gly Ala Ser Gly Leu Thr Ala Thr Met Gly Gly Ala Asp
1370
gga ggg aag gtc tgg ctg cag cct ggg gcc ctt gct cgg gcc cca 1371
Gly Gly Leu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
1420
gtg gac act ccc agc tct gaa gac ttc agg ccc aac tgc agc acc ctc 1419
Val Asp Thr Pro Ser Ser Gly Asp Phe Arg Pro Asn Cys Ser Thr Leu
1470
aac ttc acc ttg gtt ctg tca cgg aac aac ctg atc acc gtg cag cgg 1467
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
1515
gag atg ttc gcc cag ctc tgc cag ctg cgc ctg gcc acc cag 1515
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Ser His
1560
aac tgc atc tgg cag gcc aat ggc tcc cag ttc ctg cgg ctg acc 1563
Asn Cys Ile Ser Phe Ala Val Asn Gly Ser Gln Phe Leu Leu Thr
1610
ggt ctg cag gtc cta gcc ctg tcc cag aat aag ctg acc ctc tac cag 1611
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
1660
gag cac tca ttc acg gag cta cca cga ctg gag gcc ctg gac ctc gcc 1659

37

810 gct cct cgt ggc cgt cgt ccc atg cgt cct cct tgc ggc tgg 820 825
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp 2523
830 835
5 gac ctc tgg tac tgc ttc ctc cgt ggc tgg cct cct tgg cgg 2571
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850 855
ggg cgg caa agt ggg cga gat ggt ggc cct cgt cct tgc ggc ttc 2619
Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
860 865 870
10 gtc gtc ttc gac aac acg cag agc gca gtc gca gac tgg gtc tac aac 2667
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn
875 880 885
gag ctt cgg ggg cgg cgt ggc ggc tgc cgt ggc tgg gca ctc cgc 2715
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg
890 895 900 905
15 cgt cgt cgt gag gaa cgc gac tgg cgt cct ggc aac acc ctc ttt gag 2763
Leu Cys Leu Gly Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu
910 915 920
20 aac cgt tgg ggc cca gtc tat ggc agc agc agc cgt ttt ggc cgt 2811
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu
925 930 935
gac cac acg cag cgg gtc agt gtc ctc tgg ggc ggc ttc cgt cgt 2859
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu
940 945 950 955
25 ggc cag cag cgt cgt cgt ggc gac cgc aag gac gtc gtc gtc cgt cgt 2907
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Val
955 960 965
30 atc cgt agc cct gac ggc cgc ttc cgc tat gtc cgt cgt cgt cgt 2955
11a Leu Ser Pro Asp Gly Arg Ser Arg Tyr Val Arg Leu Arg Gln
970 975 980 985
35 ggc ctc tgc cgc cag agt gtc ctc ctc tgg ccc cag cag ccc agt ggc 3003
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly
990 995 1000
cag cgc agc ttc tgg ggc cag cgt ggc atg ggc cgt acc agc gac aac 3051
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn
1005 1010 1015
cac cac ttc tat aac cgg aac ttc ggc cag gga ccc agc ggc gaa 3096
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu
1020 1025 1030

45 SeqId.No.4
Met Gly Phe Cys Arg Ser Ala Leu His
5

Pro Leu Ser Leu Leu Val Gln Ala 11a Met Leu Ala Met Thr Leu Ala
10 20 25
50 Leu Gly Thr Leu Pro Ala Phe Leu Pro Cys Glu Leu Gln Pro His Gly
30 35 40
Leu Val Asn Cys Asn Trp Leu Phe Leu Lys Ser Val Pro His Phe Ser

36

540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000 10005 10010 10015 10020 10025 10030 10035 10040 10045 10050 10055 10060 10065 10070 10075 10080 10085 10090 10095 10100 10105 10110 10115 10120 10125 10130 10135 10140 10145 10150 10155 10160 10165 10170 10175 10180 10185 10190 10195 10200 10205 10210 10215 10220 10225 10230 10235 10240 10245 10250 10255 10260 10265 10270 10275 10280 10285 10290 10295 10300 10305 10310 10315 10320 10325 10330 10335 10340 10345 10350 10355 10360 10365 10370 10375 10380 10385 10390 10395 10400 10405 10410 10415 10420 10425 10430 10435 10440 10445 10450 10455 10460 10465 10470 10475 10480 10485 10490 10495 10500 10505 10510 10515 10520 10525 10530 10535 10540 10545 10550 10555 10560 10565 10570 10575 10580 10585 10590 10595 10600 10605 10610 10615 10620 10625 10630 10635 10640 10645 10650 10655 10660 10665 10670 10675 10680 10685 10690 10695 10700 10705 10710 10715 10720 10725 10730 10735 10740 10745 10750 10755 10760 10765 10770 10775 10780 10785 10790 10795 10800 10805 10810 10815 10820 10825 10830 10835 10840 10845 10850 10855 10860 10865 10870 10875 10880 10885 10890 10895 10900 10905 10910 10915 10920 10925 10930 10935 10940 10945 10950 10955 10960 10965 10970 10975 10980 10985 10990 10995 11000 11005 11010 11015 11020 11025 11030 11035 11040 11045 11050 11055 11060 11065 11070 11075 11080 11085 11090 11095 11100 11105 11110 11115 11120 11125 11130 11135 11140 11145 11150 11155 11160 11165 11170 11175 11180 11185 11190 11195 11200 11205 11210 11215 11220 11225 11230 11235 11240 11245 11250 11255 1

38

45
Met Ala Ala Pro Arg Gly Asn Val Thr Ser Leu Ser Leu Ser Ser Asn
60 65 70
Arg Ile His His Leu His Asp Ser Asp Phe Ala His Leu Pro Ser Leu
75 80 85
Arg His Leu Asn Leu Lys Trp Asn Cys Pro Pro Val Gly Leu Ser Pro
90 95 100
Met His Phe Pro Cys His Met Thr Ile Glu Pro Ser Thr Phe Leu Ala
110 115 120
Val Pro Thr Leu Glu Glu Leu Asn Leu Ser Tyr Asn Asn Ile Met Thr
125 130 135
Val Pro Ala Leu Pro Lys Ser Leu Ile Ser Leu Ser Leu Ser His Thr
140 145 150
Asn Ile Leu Met Leu Asp Ser Ala Ser Leu Ala Gly Leu His Ala Leu
155 160 165
Arg Phe Leu Phe Met Asp Gly Asn Cys Tyr Tyr Lys Asn Pro Cys Arg
170 175 180
Gln Ala Leu Glu Val Ala Pro Gly Ala Leu Leu Gly Leu Gly Asn Leu
190 195 200
Thr His Leu Ser Leu Lys Tyr Asn Asn Leu Thr Val Val Pro Arg Asn
202 210 215
Leu Pro Ser Ser Leu Glu Tyr Leu Leu Leu Ser Tyr Asn Arg Ile Val
220 225 230
Lys Leu Ala Pro Glu Asp Leu Ala Asn Leu Thr Ala Leu Arg Val Leu
235 240 245
Asp Val Gly Gly Asn Cys Arg Arg Cys Asp His Ala Pro Asn Pro Cys
250 255 260
Met Glu Cys Pro Arg His Phe Pro Gln Leu His Pro Asp Thr Phe Ser
270 275
His Leu Ser Arg Leu Glu Gly Leu Val Leu Lys Asp Ser Ser Leu Ser
285 290 295
Trp Leu Asn Ala Ser Trp Phe Arg Gly Leu Gly Asn Leu Val Leu
300 305 310
Asp Leu Ser Glu Asn Phe Leu Tyr Lys Cys Ile Thr Lys Thr Lys Ala
315 320 325
Phe Gln Gly Leu Thr Gln Leu Arg Lys Leu Asn Leu Ser Phe Asn Tyr
330 335 340
Gln Lys Arg Val Ser Phe Ala His Leu Ser Leu Ala Pro Ser Phe Gly
350 355 360
Ser Leu Val Ala Leu Lys Glu Leu Asp Met His Gly Ile Phe Phe Arg
365 370 375
Ser Leu Asp Glu Thr Thr Leu Arg Pro Leu Ala Arg Leu Pro Met Leu
380 385 390
Gln Thr Leu Arg Leu Gln Met Asn Phe Ile Asn Gln Ala Gln Leu Gly
395 400 405
Ile Phe Arg Ala Phe Pro Gly Leu Arg Tyr Val Asp Leu Ser Asp Asn
410 415 420
Arg Ile Ser Gly Ala Ser Glu Leu Thr Ala Thr Met Gly Glu Ala Asp
430 435 440
Gly Gly Glu Lys Val Trp Leu Gln Pro Gly Asp Leu Ala Pro Ala Pro
445 450 455
Val Asp Thr Pro Ser Ser Glu Asp Phe Arg Pro Asn Cys Ser Thr Leu

39

460 465 470
Asn Phe Thr Leu Asp Leu Ser Arg Asn Asn Leu Val Thr Val Gln Pro
475 480 485
Glu Met Phe Ala Gln Leu Ser His Leu Gln Cys Leu Arg Leu Ser His
490 495 500
Asn Cys Ile Ser Gln Ala Val Asn Gly Ser Gln Phe Leu Pro Leu Thr
510 515 520
Gly Leu Gln Val Leu Asp Leu Ser His Asn Lys Leu Asp Leu Tyr His
525 530 535
Glu His Ser Phe Thr Glu Leu Pro Arg Leu Glu Ala Leu Asp Leu Ser
540 545 550
Tyr Asn Ser Gln Pro Phe Gly Met Gln Gly Val Gly His Asn Phe Ser
555 560 565
Phe Val Ala His Leu Arg Thr Leu Arg His Leu Ser Leu Ala His Asn
570 575 580
Asn Ile His Ser Gln Val Ser Gln Gln Leu Cys Ser Thr Ser Leu Arg
590 595
Ala Leu Asp Phe Ser Gly Asn Ala Leu Gly His Met Trp Ala Glu Gly
605 610 615
Asp Leu Tyr Leu His Phe Phe Gln Gly Leu Ser Gly Leu Ile Trp Leu
620 625 630
Asp Leu Ser Gln Asn Arg Leu His Thr Leu Leu Pro Gln Thr Leu Arg
635 640 645
Asn Leu Pro Lys Ser Leu Gln Val Leu Arg Leu Arg Asp Asn Tyr Leu
650 655 660
Ala Phe Phe Lys Trp Trp Ser Leu His Phe Leu Pro Lys Leu Glu Val
670 675
Leu Asp Leu Ala Gly Asn Gln Leu Lys Ala Leu Thr Asn Gly Ser Leu
685 690 695
Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser Ile
700 705 710
Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg Glu
715 720 725
Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp Phe
730 735 740
Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn Pro
750 755 760
Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Val
765 770 775
Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser Pro
780 785 790
Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Cys
795 800 805
Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu Ala
810 815 820
Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly Trp
830 835 840
Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp Arg
845 850 855
Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala Phe
860 865 870
Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr Asn

40

875
Glu Leu Arg Gly Gln Leu Glu Cys Arg Gly Arg Trp Ala Leu Arg 885
890
Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe Glu 905
910
Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val Leu 920
925
Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu Leu 935
940
Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Leu Leu Val 950
955
Ile Leu Ser Pro Asp Gly Arg Arg Ser Arg Tyr Val Arg Leu Arg Gln 965
970
Arg Leu Cys Arg Gln Ser Val Leu Leu Trp Pro His Gln Pro Ser Gly 985
990
Gln Arg Ser Phe Trp Ala Gln Leu Gly Met Ala Leu Thr Arg Asp Asn 1000
1005
His His Phe Tyr Asn Arg Asn Phe Cys Gln Gly Pro Thr Ala Glu 1015
1020 1025 1030

5

10

15

20

41

CLAIMS

1. An isolated Toll-like-receptor polypeptide consisting essentially of
 - (i) the amino acid sequence SEQ ID NO: 2;
 - (ii) a variant thereof which has immunomodulatory activity; or
 - (iii) a fragment of (i) or (ii) which has immunomodulatory activity.
2. A polypeptide according to claim 1 wherein the variant (ii) has at least 70% identity to the amino acid sequence of SEQ ID NO: 2.
3. A polypeptide according to claim 2 which has at least 70% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 70% of the full-length sequence provided by SEQ ID No.1 and exhibits toll-like receptor functionality.
4. A polypeptide according to claim 1 or claim 2 wherein the variant (ii) has at least 95% identity to the amino acid sequence of SEQ ID NO: 2.
5. A polypeptide according to claim 2 which has at least 95% identity to the amino acid sequence of SEQ ID NO: 2 over a region of at least 60 contiguous amino acids and exhibits toll-like receptor functionality.
6. A polypeptide according to claim 1 wherein the fragment (iii) is a peptide of up to 150 amino acids in length and exhibits toll-like receptor functionality.
7. A polynucleotide encoding a polypeptide according to any one of claims 1-3.
8. A polynucleotide encoding a Toll-like receptor polypeptide which has immunomodulatory activity, which polynucleotide consists essentially of:
 - (a) the nucleic acid sequence of SEQ ID NO: 1;
 - (b) a sequence complementary thereto;

5

10

15

20

25

30

42

- (c) a sequence which hybridises under stringent conditions to a sequence as defined in (a) or (b);
 (d) a sequence that is degenerate as a result of the genetic code to a sequence as defined in (a), (b) or (c); or
 (e) a sequence having at least 60% identity to a sequence as defined in (a), (b), (c) or (d).

9. A polynucleotide according to claim 7 or claim 8 which is mRNA.

10. A polynucleotide according to claim 7 or claim 8 which is DNA.

11. A polynucleotide according to claim 7 or claim 8 which is cDNA.

12. An isolated toll-like receptor polypeptide which is obtainable by expression *in vivo* or *in vitro* of a polynucleotide according to claim 7 or claim 8.

13. A polypeptide according to claim 12 which has the structural features conserved amongst toll-like receptors.

14. An expression vector comprising a polynucleotide sequence according to any one of claims 7 to 11, which is capable of expressing a polypeptide according to any one of claims 1 to 3 or claim 12.

15. An expression vector according to claim 14 which is a plasmid, phage or viral vector.

16. A host cell comprising an expression vector according to claim 14 or claim 15.

43

17. A polyclonal or monoclonal antibody, or a chimera or fragment thereof, which is specific for a polypeptide according to any one of claims 1 to 3.

18. A method for identification of a compound that modulates Toll-like receptor activity, which method comprises contacting a polypeptide according to any one of claims 1 to 3 or claim 12 with a test substance and monitoring for immunomodulatory activity.

19. A compound which modulates Toll-like receptor activity and which is identifiable by a method according to claim 18.

20. A compound according to claim 19 which is a peptide or polypeptide.

21. A compound according to claim 19 which is an oligonucleotide or polynucleotide.

22. A method of treating a subject having an inflammatory or cardiovascular disorder, systemic infection or autoimmune disease that is responsive to Toll-like receptor modulation, which method comprises administering to said subject an effective amount of a compound according to any one of claims 19 to 21 or an antibody according to claim 17.

23. A method according to claim 22 wherein the disorder is a viral, fungal or bacterial infection, asthma, rhinitis, chronic obstructive pulmonary disease (COPD), emphysema, an inflammatory bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriasis, Alzheimer's disease, atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome associated with systemic infection involving gram positive or gram negative bacteria.

24. A polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, for use in the treatment of an immune or inflammatory disorder.

25. A polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 for use as an adjuvant.

26. The use of a compound according to claim 19 in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

27. The use of a polypeptide comprising a fragment or variant of SEQ ID NO: 2, which is capable of inhibiting the activity of TLR9 having the amino acid sequence of SEQ ID NO: 2 or a functional variant thereof, in the manufacture of a medicament for the treatment of an immune or inflammatory disorder.

28. The use of a polypeptide according to any one of claims 1 to 3 or claim 12, a polynucleotide according to claim 7 or claim 8 or a compound according to claim 19 in the manufacture of an adjuvant or vaccine formulation.

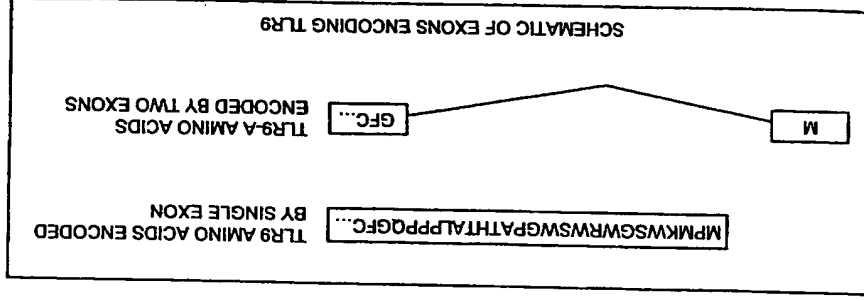
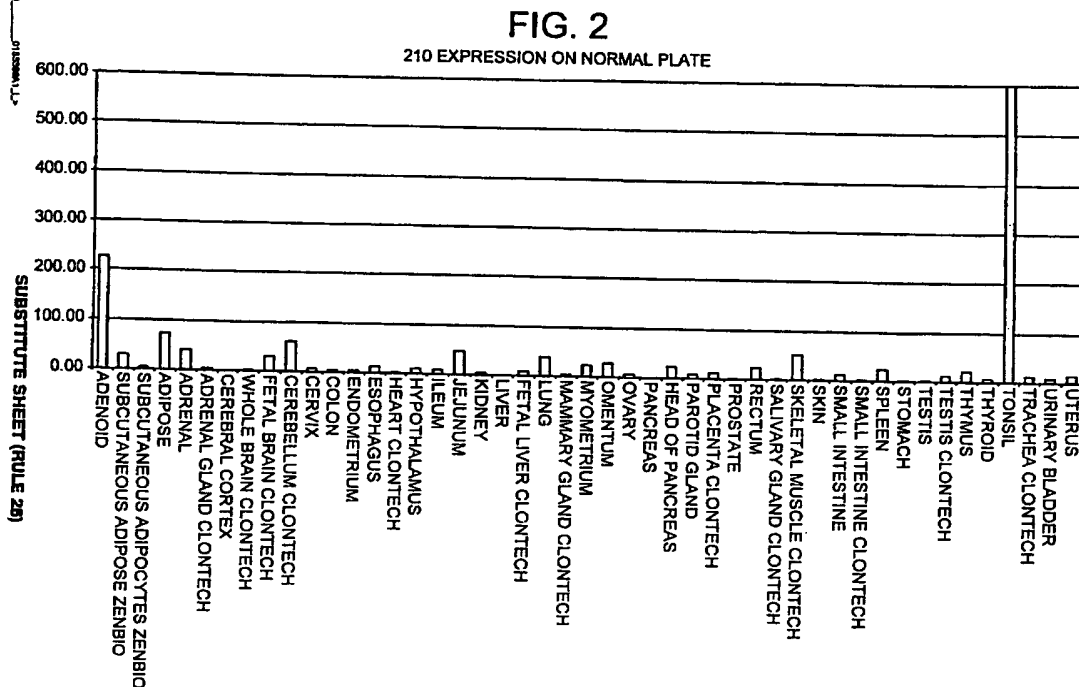


FIG. 1

2 / 5



3 / 5

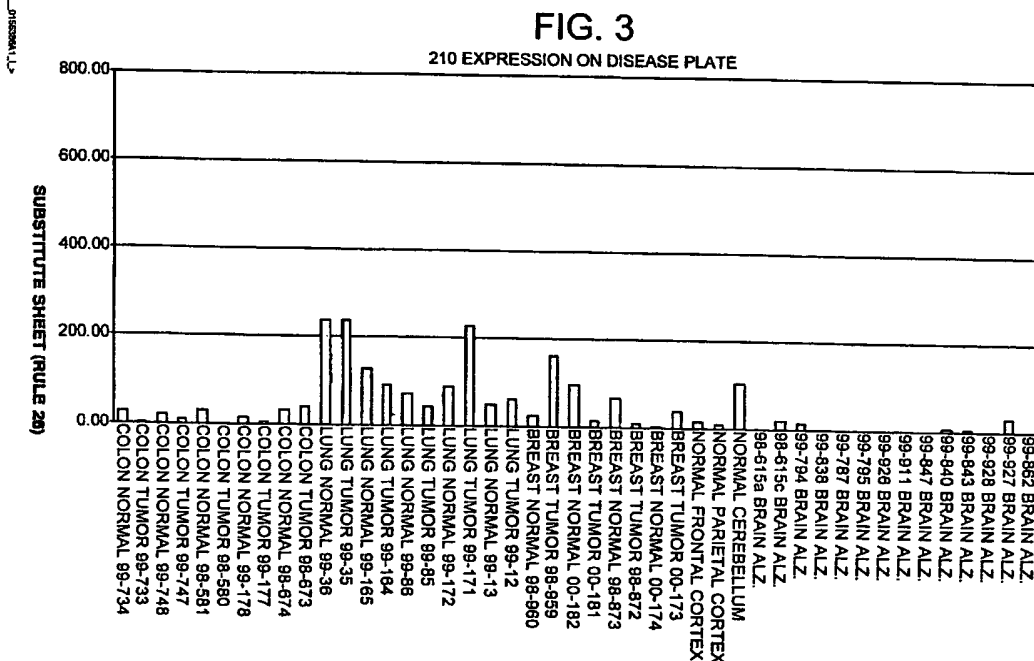


FIG. 3CONT'D

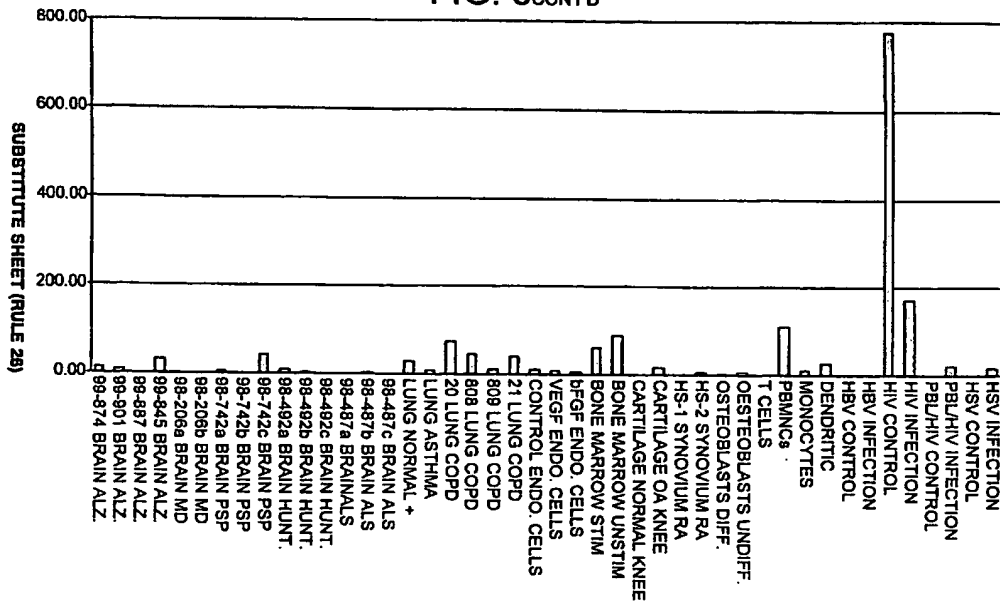
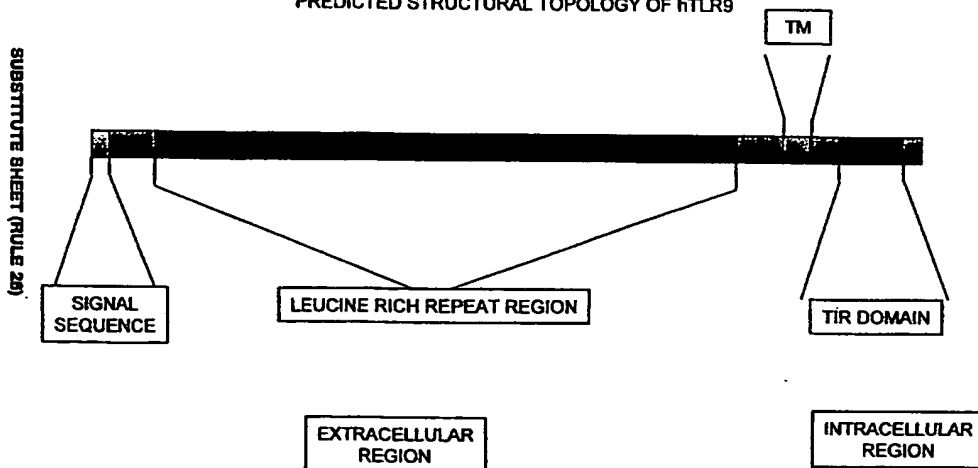


FIG. 4

PREDICTED STRUCTURAL TOPOLOGY OF hTLR9



INTERNATIONAL SEARCH REPORT

<p>Intem. Appl. Application No. PCT/GB 01/00299</p>	
<p>A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/63 C12N5/10 G01N33/68 C07K14/705 C12N16/28 A61K31/7088 A61K38/17 A61K48/00</p>	
<p>According to International Patent Classification (IPC) or to both national classification and IPC</p>	
<p>B. FIELD SEARCHED IPC 7 C12N G01N C07K A61K</p>	
<p>Documents searched other than mentioned in the above that such documents are included in the fields searched</p>	
<p>Electronic data base consulted during the international search (name of data base and, where practical, search terms used) ENBL, STRAND, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, CHEM ABS Data</p>	
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p>	
<p>Category 1 Citation of document, with indication, where appropriate, of the relevant passages</p>	<p>Relevant to claim No.</p>
<p>X NO 98 50547 A (SCHERING CORP) 12 November 1998 (1998-11-12) 99.8% identity in 1008 bp overlap between SEQ ID NO 33 of W09850547 and SEQ ID NO 1 99.7% identity in 336 amino acids overlap between SEQ ID NO 34 of W09850547 and SEQ ID NO 2 page 56, line 23 -page 60, line 14: claims 1-16; examples 1-11</p>	<p>1-18; 22-25, 27, 28</p>
<p>Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.</p>	
<p>* Special categories of cited documents:</p>	
<p>*V document defining the general state of the art which is not considered to be of particular relevance</p>	<p>*T later documents published after the international filing date of the claimed invention which are not considered to be of particular relevance</p>
<p>*W document published on or after the international filing date which may throw doubt on priority claim(s) or which is cited to establish the publication date of another document</p>	<p>*X document of particular relevance the claimed invention involves an inventive step when the document is taken alone or in combination with other prior art</p>
<p>*Y document published in an official gazette, journal, or other means</p>	<p>*Z document published prior to the international filing date but after the priority date claimed</p>
<p>Date of the actual completion of the international search</p>	<p>Date of mailing of the international search report</p>
<p>28 June 2001</p>	<p>11/07/2001</p>
<p>Name and mailing address of the ISA European Patent Office, P.O. Box 1618, 6018 Paderborn 2 Tel. (+31-70) 340-0540, Tel. 31 651 650 01 Fax (+31-70) 340-0518</p>	<p>Authorized officer Devijver, K</p>

INTERNATIONAL SEARCH REPORT

<p>Intem. Appl. Application No. PCT/GB 01/00299</p>	
<p>C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT</p>	
<p>Category 1 Citation of document, with indication, where appropriate, of the relevant passages</p>	<p>Relevant to claim No.</p>
<p>X DATABASE EMBL 'Online' accession: AC006252 29 December 1998 (1998-12-29) MUZY D ET AL: "Homo sapiens 3p21.1 contig 9 PAC RPC15-1157M23 (Roswell Park Cancer Institute Human PAC Library) complete sequence." XP002170407 100% identity in 3165 bp overlap (between positions 26803-29967) with SEQ ID NO 1 100% identity in 1055 amino acids overlap (between positions 26803-29967) with SEQ ID NO 2 (tfasta)</p>	<p>7-11</p>
<p>P, X HEMMI HIROAKI ET AL: "A Toll-like receptor recognizes bacterial DNA" NATURE, MACMILLAN JOURNALS LTD, LONDON, GB, vol. 408, no. 6813, 7 December 2000 (2000-12-07), pages 740-745, XP002168474 ISSN: 0028-0836 cited in the application the whole document -8 DATABASE EMBL 'Online' accession: AB045180, 13 December 2000 (2000-12-13) AKIRA S ET AL: "Homo sapiens TLR9 mRNA for Toll-like receptor 9, complete cds." XP002170408 99.7% identity in 3113 bp overlap with SEQ ID NO 1 100% identity in 1031 amino acids overlap with SEQ ID NO 2</p>	<p>1-18, 22-25, 27, 28</p>

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report		Publication date		Patent family member(s)		Publication date	
NO 9850547	A	12-11-1998	A	AU	7175498	A	27-11-1998
				BR	9808747	A	11-07-2000
				CN	1263555	T	16-08-2000
				EP	0980429	A	23-02-2000
				NO	995458	A	08-11-1999
				PL	336635	A	03-07-2000
				SK	146599	A	11-07-2000
				HU	0001462	A	28-07-2000

International Application No. PCT/88 01 00299

FURTHER INFORMATION CONTINUED FROM PCT/88A/ 210

Continuation of Box I.2

Claims Nos.: 19-21,26; in part: 22,23,25,28

Claim 19 refers to a compound which modulates Toll-like receptor activity without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claim is ambiguous and vague, and its subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).

No meaningful search can be carried out for such a purely speculative claim whose wording is, in fact, a mere recitation of the result to be achieved.

The above comment also applies for claims 20,21 and 26; and in part for claims 22,23,25 and 28.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

From PCT/88A/210 (parent family applied 1 July 1988)

INSDOCID: -WFO_01E388A1.JJ>

INSDOCID: -WFO_01E388A1.JJ>

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)